417.6
RA1
1
c1

ラ　オ

統計的推測とその応用

奥野忠一・長田　洋
篠崎信雄・広崎昭太
古河陽子・矢島敬二
鷲尾泰俊　　訳

東京図書株式会社

序

この本の目的は，統計的推測の最新の理論とその諸手法を体系的に整理し，かつ実用向きに説明することである．この主題について，最近30年間に得られた重要な成果の要点はすべて取り入れた．読者には，大学学部程度の数学と統計学の基礎知識を期待している．

定量的な推測に科学的な性格をもたせるべきであるなら，その論理的・数学的・確率論的視点をはずすことはできない．統計的推測に至る主要な道筋は帰納的推理であり，これによって“不確定さを含む陳述”が導かれる．不確定さの程度を厳密に表現するためには，現代統計学の基礎を成す数学的方法と確率の諸概念が必要である．統計的推測についてそれだけで完備した議論を行うには，高度の数学的方法と確率論の説明が絶対必要であるとかねて考えていたので，本書の第1章ではベクトル空間と行列の方法をくわしく扱い，第2章は確率の測度論的表現と確率の道具と手法の展開にあてることにした．

統計的推測の諸手法は，現実世界に適用されなければその意味を失い，演繹法の練習のようになってしまう．その上，統計コースにおいては，統計学の数学理論とその理論の実際問題への適用の両方に重点をおかねばならないというのが，私の信念である．統計的手法の適用についてくわしく論議すれば，その手法の背後にある理論がよりよく理解されるようになる．このために，本書では，生きた例題を数学的な結果と織りまぜて述べる．そのほかに，各章の終わりに多数の問題を与えた．その一部は，その章の本文中に導いた主要な結果を補足するためのものであり，他は，理論についての読者の理解度をためすための練習問題用のものである．

統計的推測の広い分野をすべて包含するように題材を選んで説明することは容易ではなかった．だが，過去20年間の，学部および大学院学生の教育と統計学研究の遂行および指導における私自身の経験が役に立った．統計学の主要な方法を選んで説明し，その理論的な基礎を詳細に論じることにより，読者が統計学について助言したり，特殊な問題の学習や研究を遂行したりするための素養を得られるようにしたつもりである．

第1章でベクトルと行列の代数をやや長すぎるくらいに述べた理由を少し説明しておきたい．第1は，あとの章での統計的手法の数学的取り扱いが，ベクトル空間と行列の方法に大きく依存しているためである．第2は，ベクトルと行列の代数は，物理系，生物系，

社会系科学の最近の論文で広く用いられるようになった数学の一部門を構成するからである．この章の主題には論理的かつ厳密な取り扱いがなされており，徐々に高度の水準へと展開されている．重要な定理とそれから導かれる諸結果はすべて，科学の諸分野の研究者も容易に利用できるような形で説明した．

第2章の内容は，統計的推測を取り扱うのに必要な確率の道具と手法を体系的にまとめたものである．確率の公理からはじめ，確率変数，分布関数，条件づき期待値と条件づき分布の諸概念の定式化に進む．それにつづいて，特性関数，無限次元積空間における確率分布，および，重要なすべての極限定理の研究が説明される．第2章ではまた，他のいくつかの章でしばしば用いられ，かつ，高度な確率論を専攻しようとする人々には良い素養として役立つような，数多くの命題が述べられている．

第3章は，統計的推測に必要な連続型確率模型と標本分布を取り扱う．正規分布，ガンマ分布，コーシー分布などのような実際にしばしば用いられるいくつかの重要な分布を，観測値を生成する物理的機構についての適切な確率模型を通じて導入する．この章の1つの特色は，粒子の平衡分布に関する統計力学の問題の論議である．

第4章は，分散分析の手法を用いる推測にあてた．ガウス＝マルコフの線形模型とこれについての推定，検定の問題を，十分一般的に取り扱う．分散成分の問題は，級内相関係数の推定という，より一般的な問題の特殊の場合とみなして考える．多重分類データで，カテゴリー内の級ごとにサンプリング方式が異なる場合の統一的な取り扱いについても述べる．

推定に関する各種の理論と手法が，第5章の主題である．推定問題についての論争の一部をここで考察する．そして，現存する不一致のある部分を除去するために，大標本における推定の基準に若干の修正をもちこんだ．

規定の問題，および，それに関連する類似標本や類似推定値の一様性の検定を，第6章で扱う．観測値がそれに従って得られたと考えることのできる数学模型の選択は基本的に重要であって，それにつづく統計計算はこの選ばれた模型のわく組のなかで行われる．入手した事実に基づいて，提案された模型の妥当性を検証するための適切な検定法が開発された．

第7章では，仮説の検定，区間推定，実験，識別問題，ノンパラメトリック推測などのような諸種の統計的推測の理論的背景を述べる．

最後の第8章では，多変量データに基づく推測を取り扱う．この章の特色は，多変量正規分布の研究を，その密度関数を用いないで，簡単な特徴づけによって行ったことである．この特徴づけによって多変量の理論は簡単になり，1変量の理論からの適切な一般化が余分の解析なしに可能になった．また，これによって，ヒルベルト空間での分布のようなより一般的な場での多変量正規分布を研究するための必要な知識が得られた．

本書全体を通じて，節などの引用を示すのに一定の記号を用いる．その読み方は次の例のとおりである．4f.3 という小節は，第4章 f 節の第3小節である．式 (4f.3.6) は 4f.3 小節の6番目の式，表 4f.3β は 4f.3 小節の2番目の表である．各小節の主要な命題（または定理）には，(i)，(ii) などの番号をつけてある．前出の［(iii)，5d.2］というような引用は，5d.2 小節の命題 (iii) を指す．

本書の主要な部分は，アメリカ国立科学財団（National Science Foundation）の上級科学者フェローシップを得て，1963—1964年に，バルチモアのジョーンズ・ホプキンス大学客員教授であったときに書いた．ジョーンズ・ホプキンス大学では，統計学教授，G. S. Watson がこの原稿を執筆の各段階で読み，絶えず助言を与えてくださった．第7，8章についての Herman Chernoff，第2章についての Rupert Miller と S. W. Dharmadhikari，および，第1，3章についての Ralph Bradley のコメントは，最終原稿を書くのに非常に役立った．これらの人すべてに感謝の意を表したい．原稿のタイプと印刷のための整理という厄介な仕事を非常に注意深くかつ勤勉に遂行してくれた G. M. Das の援助がなければ，この原稿の執筆と改訂は容易な仕事ではなかったであろう．

最後に，故 Ronald A. Fisher 卿と P. C. Mahalanobis 教授に，私の感謝の念を捧げたい．この2人の感化によって，私は，統計学を今世紀の新しいテクノロジーであると認識するようになったのである．

インド，カルカッタ市
1965年6月　　C. R. ラオ

第 2 版の序

この本の目的は，第 1 版と同様に，統計について助言する人々のために，広汎な統計的方法のあらゆる角度からの論議を 1 冊の本にまとめ，同時に，研究者が熟知すべき統計的手順を導くために必要な数学的，論理的道具立てを厳密な形で説明することであった．

新しい題材を数多く追加したので，理論および応用の両面におけるいくつかの点で，この本は最新のものになった．

追加した重要なものは，種々の型の一般逆行列，統計量と部分集合体の概念，分散成分の MINQUE 理論，重複対数の法則と検出力 1 の逐次検定，構造母数をもつ分散共分散分析，複合諸仮説のあいだの判別，成長モデル，特性関数の理論，などである．

特に言及したいのは，観測値が必ずしも正則ではない共分散行列をもつ場合の線形模型における母数推定に関する新しい話題についてである．Gauss (1809) および Aitken (1935) による現存の理論と方法は，共分散行列が正則であることがわかっている場合にのみ適用可能である．この本 (4 i 節) で論じる新しい統一理論は，共分散行列が正則であると否とを問わず，あらゆる場合に有効である．

新しい練習問題と補足も多数つけ加えた．

第 2 版を準備しているときに助言してくださった，Dr.M.S.Avadhani, Dr.J.K.Ghosh, Dr. A. Maitra, Dr.P.E.Nüesch, Dr. Y. R. Sarma, Dr. H. Toutenberg, Dr. E.J.Williams に謝意を表したい．

ニュー・デリーにて

C. R. ラオ

目　次

序

第 2 版の序

第 1 章　ベクトルと行列の代数 …… 1

1a　ベクトル空間 …… 2

1a. 1　ベクトル空間と部分空間の定義　1a. 2　ベクトル空間の基底　1a. 3　1 次方程式　1a. 4　内積をもつベクトル空間

補足と問題 …… 11

1b　行列と行列式の理論 …… 13

1b. 1　行列演算　1b. 2　基本行列と行列の対角化　1b. 3　行列式　1b. 4　線形変換　1b. 5　一般逆行列　1b. 6　ベクトル空間，基底などの行列表現　1b. 7　ベキ等行列　1b. 8　行列の特殊な積

補足と問題 …… 29

1c　固有値と行列の標準化 …… 32

1c. 1　2 次形式の分類と変換　1c. 2　固有方程式の根　1c. 3　行列の正準化　1c. 4　射影作用素　1c. 5　一般逆行列に関するその他の結果　1c. 6　制限づき固有値問題

1d　ベクトル空間の凸集合 …… 48

1d. 1　定義　1d. 2　凸集合の分離定理

1e　不等式 …… 51

1e. 1　コーシー＝シュヴァルツ不等式　1e. 2　ヘルダーの不等式　1e. 3　アダマールの不等式　1e. 4　積率に関する不等式　1e. 5　凸関数とイェンセンの不等式　1e. 6　情報理論における不等式　1e. 7　スターリングの近似

1f　2 次形式の極値 …… 56

1f. 1　一般的な結果　1f. 2　固有値と固有ベクトルに関連する結果　1f. 3　対角和最小問題

補足と問題 …… 63

第2章　確率論，道具と方法 …… 73

2a　確率の計算 …… 74

2a.1　基本事象の空間　2a.2　部分集合（事象）の族　2a.3　集合関数としての確率　2a.4　ボレル集合体（σ 集合体）と確率測度の拡張　2a.5　確率変数と分布関数の概念　2a.6　多次元確率変数　2a.7　条件つき確率と統計的独立性　2a.8　確率変数の条件つき分布

2b　数学的期待値と確率変数の積率 …… 86

2b.1　数学的期待値の性質　2b.2　積率　2b.3　条件つき期待値　2b.4　特性関数　2b.5　反転定理　2b.6　多変量積率

2c　極限定理 …… 100

2c.1　コルモゴロフの一致性定理　2c.2　確率変数列の収束　2c.3　大数の法則　2c.4　分布関数列の収束　2c.5　中心極限定理　2c.6　独立確率変数の和

2d　確率測度の族と統計学の問題 …… 123

2d.1　確率変数の族　2d.2　十分統計量の概念　2d.3　十分性の特徴づけ

付録2A　スティルチェス積分，ルベーグ積分 …… 125
付録2B　測度論と積分における重要な定理 …… 127
付録2C　不変性 …… 130
付録2D　統計量，部分空間，十分性 …… 131
付録2E　特性関数の非負定符号性 …… 133
補足と問題 …… 134

第3章　連続型確率模型 …… 145

3a　1変量模型 …… 147

3a.1　正規分布　3a.2　ガンマ分布　3a.3　ベータ分布　3a.4　コーシー分布　3a.5　スチューデントの t 分布　3a.6　統計力学での平衡状態をあらわす分布　3a.7　円周上の分布

3b　標本分布 …… 165

3b.1　定義と結果　3b.2　正規確率変数の2乗和　3b.3　標本平均と標本分散の同時分布　3b.4　2次形式の分布　3b.5　最小2乗法の3つの基本定理　3b.6　p 変量正規分布　3b.7　指数型分布族

3c　対称正規分布 …… 182

3c.1　定義　3c.2　標本分布

3d　2変量正規分布 …… 186

3d.1　一般的性質　3d.2　標本分布

補足と問題 …… 192

第4章　最小2乗法と分散分析の理論 …… 201

4a　最小2乗法の理論（線形推定論） …… 201

4a.1　ガウス＝マルコフ模型($\mathbf{Y}, \mathbf{X}\beta, \sigma^2\mathbf{I}$)　4a.2　正規方程式と最小2乗（l.s.）推定量　4a.3　一般逆行列と正規方程式の解　4a.4　l.s. 推定量の分散と共分散　4a.5　σ^2 の推定　4a.6　最小2乗法の理論にたいする他の接近（幾何学的解）　4a.7　観測値に相関のある場合の具体的な表現　4a.8　最小2乗法の理論における計算上の問題　4a.9　母数に制約があるときの最小2乗推定　4a.10　母数の関数の同時推定　4a.11　母数が確率変数であるときの最小2乗法の理論　4a.12　計画行列の選択

4b　仮説検定と区間推定 …… 217

4b.1　母数の1つの関数（の推測）　4b.2　母数の2つ以上の関数（の推測）　4b.3　制約のある模型

4c　単一標本の問題 …… 222

4c.1　検定基準　4c.2　左右の大腿骨の非対称性（一対比較）

4d　1元分類のデータ …… 224

4d.1　検定基準　4d.2　1つの例題

4e　2元分類のデータ …… 226

4e.1　各級で1つの観測値の場合　4e.2　各級で2つ以上の等しい数の観測値がある場合　4e.3　級における繰り返し数が等しくない場合

4f　2元データにたいする一般模型と分散成分 …… 236

4f.1　一般模型　4f.2　分散成分模型　4f.3　一般模型の処理

4g　統計的回帰の理論と応用 …… 241

4g.1　回帰の概念（一般理論）　4g.2　関連性の増加の尺度　4g.3　頭蓋容量の予測（1つの実例）　4g.4　回帰式の一様性の検定　4g.5　指定された回帰関数にたいする検定　4g.6　制約のある回帰

4h　2組の母数をもつ最小2乗法の一般的問題 …… 262

4h.1　補助変数　4h.2　共分散分析　4h.3　1つの例題

4i　線形推定の統一理論 …… 267

4i.1　一般逆行列に関する基本的補助定理　4i.2　一般ガウス＝マル

コフ模型　4i.3　逆分割行列法　4i.4　最小2乗法の統一理論

4j　分散成分の推定 …… 275

4j.1　分散成分模型　4j.2　MINQUE 理論　4j.3　ユークリッド・ノルムのもとでの計算

4k　線形模型における偏りのある推定 …… 278

4k.1　最良線形推定 (BLE)　4k.2　最良線形偏り最小推定 (BLIMBE)

補足と問題 …… 281

第5章　推定の規準と方法 …… 287

5a　最小分散不偏推定 …… 287

5a.1　最小分散規準　5a.2　最小分散推定についてのいくつかの基本的な結果　5a.3　母数が複数個ある場合　5a.4　フィッシャーの情報の尺度　5a.5　不偏推定量の改良

5b　一般的手順 …… 305

5b.1　一般的問題の記述（ベイズの定理）　5b.2　$(\theta, \mathbf{X})$ の同時分布関数が完全に知られている場合　5b.3　同程度に未知とみなす法則　5b.4　経験的ベイス推定の手順　5b.5　推測確率　5b.6　ミニマックス原理　5b.7　不変性の原理

5c　大標本における推定の規準 …… 315

5c.1　一致性　5c.2　有効性

5d　大標本における推定の方法 …… 321

5d.1　モーメント法　5d.2　最小カイ2乗法およびそれに関連した方法　5d.3　最尤法

5e　多項分布の推定 …… 326

5e.1　確率が母数に依存しない場合　5e.2　確率が母数の関数である場合

5f　一般の場合における母数の推定 …… 333

5f.1　仮定と記号　5f.2　m.l. 方程式推定量の性質

5g　母数の推定における得点による方法 …… 336

補足と問題 …… 342

第6章　大標本の理論と方法 …… 349

6a　基本的な諸結果 …… 349

6a.1　度数の2次関数の漸近分布　6a.2　収束定理

6b　多項分布のカイ2乗検定 …… 357

6b.1　単純仮説からのずれの検定　6b.2　適合度についてのカイ2乗検定　6b.3　1つの級における偏差の検定　6b.4　母数がある部分集合に含まれるか否かの検定　6b.5　例題　6b.6　いくつかの級における偏差の検定

6c　多項分布からの独立標本に関する検定 …… 363

6c.1　一般的な結果　6c.2　類似標本の一様性の検定　6c.3　例題

6d　分割表 …… 367

6d.1　観測結果の得られる確率と大標本の検定　6d.2　分割表における独立性の検定　6d.3　小標本における独立性の検定

6e　いくつかの一般的な大標本検定 …… 378

6e.1　記号と基本的な結果　6e.2　単純仮説の検定　6e.3　複合仮説の検定

6f　順序統計量 …… 382

6f.1　経験分布関数　6f.2　標本分位点の漸近分布

6g　統計量の変換 …… 388

6g.1　一般公式　6g.2　ポアソン変量の平方根変換　6g.3　2項割合の平方根の逆正弦変換　6g.4　相関係数の $\tanh^{-1}$ 変換

6h　積率の標準誤差と関連統計量 …… 396

6h.1　原点まわりの積率の分散と共分散　6h.2　平均値まわりの積率の漸近分散と漸近共分散　6h.3　平均値まわりの積率の分散・共分散の正確な式

補足と問題 …… 399

第7章　統計的推測の理論 …… 403

7a　統計的仮説検定 …… 403

7a.1　問題の記述　7a.2　ネイマン＝ピアソンの基本補助定理と一般化　7a.3　単純仮説 H にたいする単純仮説 H_0　7a.4　局所最強力検定　7a.5　複合仮説の検定　7a.6　フィッシャー＝ベーレンス問題　7a.7　検定の漸近効率

7b　信頼区間 …… 426

7b.1　一般的な問題　7b.2　信頼集合を構成する一般的な方法　7b.3　θ の関数の集合推定量

7c 逐次解析 …… 429
7c. 1 ワルドの逐次確率比検定 (S.P.R.T.) 7c. 2 S. P. R. T. のいくつかの性質 7c. 3 S. P. R. T. の効率 7c. 4 逐次検定の経済性の例 7c. 5 逐次解析の基本等式 7c. 6 逐次推定
7c. 7 検出力が 1 の逐次検定

7d 識別問題——決定理論 …… 444
7d. 1 問題の記述 7d. 2 確率化決定方式と非確率化決定方式
7d. 3 ベイズ解 7d. 4 決定方式の完全類 7d. 5 ミニマックス方式

7e ノンパラメトリック推測 …… 450
7e. 1 頑健性の概念 7e. 2 分布によらない方法 7e. 3 いくつかのノンパラメトリック検定 7e. 4 無作為化の原理

7f 補助情報 …… 457

補足と問題 …… 458

第 8 章 多変量解析 …… 467

8a 多変量正規分布 …… 468
8a. 1 定義 8a. 2 分布の性質 8a. 3 N_p のいくつかの特徴づけ 8a. 4 多変量正規分布の密度関数 8a. 5 母数の推定
8a. 6 最大エントロピーをもつ分布としての N_p

8b ウィシャート分布 …… 482
8b. 1 定義と記号 8b. 2 ウィシャート分布に関する若干の結果

8c 分散共分散分析 …… 491
8c. 1 多重測定にたいするガウス＝マルコフ模型 8c. 2 母数の推定
8c. 3 線形仮説の検定，分散共分散分析 8c. 4 付加情報にたいする検定 8c. 5 Λ の分布 8c. 6 次元数の検定（構造関係）
8c. 7 構造母数をもつときの分散共分散分析（成長モデル）

8d 多変量検定のいくつかの適用 …… 509
8d. 1 指定された平均値の検定 8d. 2 平均値の所与の構造にたいする検定 8d. 3 2 つの母集団の平均値の差の検定 8d. 4 数個の母集団の間の平均値の差の検定 8d. 5 頭蓋骨特性における永年変異に関するバーナードの問題

8e 判別分析（識別問題） …… 518
8e. 1 決定のための判別得点 8e. 2 研究における判別分析
8e. 3 複合仮説間の判別

8f 変量の組のあいだの関係 …… 527
8f. 1 正準相関 8f. 2 正準変数のもつ性質 8f. 3 共通因子の有効個数 8f. 4 因子分析

8g 確率変数の正規直交基底 …… 531
8g. 1 グラム＝シュミット基底 8g. 2 主成分分析

補足と問題 …… 536

著者の刊行物 …… 547

索引 …… 555

訳者あとがき …… 565

第1章　ベクトルと行列の代数

まえがき　行列の理論は，純粋数学でも，自然科学，社会科学でも，現在では広く利用されている．しかしながら，ベクトル空間と線形変換（行列はその特殊な場合である）の理論は，より基礎的であり問題のより良い理解に役立つにもかかわらず，これまであまり重要な地位を占めてこなかった．ベクトル空間の概念は，ゲーム，経済行動，時系列予測の理論や，1変量，多変量統計的方法の最新の取り扱いなどの話題を論ずるのに欠くことのできないものである．

第1章の目的は，ベクトル空間の概念とそれから導かれる基本的な結果を読者に紹介することである．重要な定理はすべて十分詳細に論議し，初心者でもこの章を完全に読了できるようにしてある．この主題の理解を一層深めるために，多数の例題と問題をあげておいた．広く一般的に論ずるために（このことは重要ではあるが，それによって理論の理解を困難にしてはならない），ベクトルの演算に用いる要素は，加法 addition と乗法 multiplication が（実数系におけるように）矛盾なく定義される体 Field に属すると考える．つまり，要素 $e_1, e_2, \cdots$（その数は有限でも無限でもよい）が体Fに属するとは，それらが加法 $e_i + e_j$ と乗法 e_ie_j に関して閉じている場合，すなわち，Fの要素の和と積がまたFに属し，かつ，次の条件を満足する場合をいう：

(A_1)　$e_i + e_j = e_j + e_i$　（交換律）

(A_2)　$e_i + (e_j + e_k) = (e_i + e_j) + e_k$　（結合律）

(A_3)　任意の2つの要素 e_i, e_j にたいして，$e_i + e_k = e_j$ となるような要素 e_k が存在する．

条件 (A_3) から，すべての i について，$e_i + e_0 = e_i$ を満足するような要素 e_0 が存在することがわかる．この要素 e_0 は，実数系における0（ゼロ）と同様のものである．

(M_1)　$e_ie_j = e_je_i$　（交換律）

(M_2)　$e_i(e_je_k) = (e_ie_j)e_k$　（結合律）

(M_3)　$e_i(e_j + e_k) = e_ie_j + e_ie_k$　（分配律）

(M_4)　任意の2つの要素 e_i, e_j にたいして，$e_i \neq e_0$（ゼロ要素）ならば，$e_ie_k = e_j$ となるような要素 e_k が存在する．

条件 (M_4) から，すべての i について，$e_ie_1 = e_i$ を満足するような要素 e_1 が存在する

ことがわかる．この要素 e_1 は，実数系における1のようなものである．

ベクトル空間の勉強をした後で，最新の行列理論と2次形式の議論をする．その基本的な命題のほかに，数理物理学，経済学，生物学，統計学で用いられる多くの結果と数値計算法をまとめて，統一した形で示す．これは，自然科学，社会科学への行列理論の適用に関心のある人々の役に立つであろう．

1a　ベクトル空間

1a.1　ベクトル空間と部分空間の定義

力，生物体の大きさ，個人の健康とか知能，物価水準などの概念は，ただ1つの数値で完全にあらわすことはできない．それらは，いろいろな方向への現れ方によって理解しなければならず，その各方向への現れ方は1つの数値であらわしてもよい．個人の知能は，特定の k 種のテストにおける彼の得点 $(x_1, x_2, \cdots, x_k)$ で判断できるであろう．このように順序づけた測定値の組を，単に $\mathbf{x}$ であらわし，行ベクトルと呼ぶ．いま，$\mathbf{y} = (y_1, y_2, \cdots, y_k)$ を別の人の得点ベクトルとすると，これらのテストにおける2人の合計点は，

$$\mathbf{x} + \mathbf{y} = (x_1 + y_1, x_2 + y_2, \cdots, x_k + y_k) \qquad (1a.1.1)$$

によって定義されるベクトル $\mathbf{x} + \mathbf{y}$ によってあらわされる．2つのベクトルを結合する，つまり加えるときの規則は，2次元または3次元における2つの力を合成するときに周知の，**力の平行四辺形則**と同じである．代数的には，この規則は，その成分が個々の力の対応する成分の和であるような力を求めることと同等である．

力 $\mathbf{f} = (f_1, f_2, f_3)$ が与えられたとき，その c 倍の力として，

$$c\mathbf{f} = (cf_1, cf_2, cf_3) \qquad (1a.1.2)$$

を定義するのは自然であり，これからベクトルに c のようなスカラーをかける新しい演算が導かれる．さらに，力 $\mathbf{f}$ が所与のとき，これに力 $\mathbf{g} = (-f_1, -f_2, -f_3) = (-1)\mathbf{f}$ を加える，つまり反対方向から力 $\mathbf{f}$ を加えることによって相殺することができ，その結果の力は $\mathbf{f} + \mathbf{g} = (0,0,0)$ となる．こうして，$-\mathbf{f}$ のような負のベクトル，および零ベクトル $\mathbf{0} = (0,0,0)$ の概念が得られる．この後者は，すべての $\mathbf{x}$ にたいして $\mathbf{x} + \mathbf{0} = \mathbf{x}$ という性質をもつ．

この新しい量〔ベクトル〕は，上で定義した演算の許す範囲で，ふつうの数と同様に取り扱えることがわかる．実際，順序づけられたスカラーの集合というような特殊の型の新しい量にわれわれの代数を制限する必要はなく，ベクトルと呼ぶことにする有限または無限個の要素 $\mathbf{x}, \mathbf{y}, \mathbf{z}, \cdots$ と，（実数系のように，適当に定義された加法，減法，乗法，除法の演算のできる）体を構成するスカラー $c_1, c_2, \cdots$ との集合を考え，それらを結合する若干の規則を定めればよい．

ベクトルの加法　＋で示される加法は，その集合の任意の2つのベクトルから1つのベクトルを導く演算として定義され，次の規則が仮定される：

（i）$\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$　（交換律）

（ii）$\mathbf{x} + (\mathbf{y} + \mathbf{z}) = (\mathbf{x} + \mathbf{y}) + \mathbf{z}.$　（結合律）

零要素　その集合には，$\mathbf{0}$ であらわされ，

（iii）$\mathbf{x} + \mathbf{0} = \mathbf{x}$,　すべての $\mathbf{x}$ にたいして

を満足する要素が存在する．

逆(負)要素　任意の要素 $\mathbf{x}$ に対応して，

（iv）$\mathbf{x} + \xi = \mathbf{0}$

を満足する要素 ξ が存在する．

スカラー倍　ベクトル $\mathbf{x}$ とスカラー c との積は，その集合の1つのベクトルとなり，$c\mathbf{x}$ であらわし，次の規則に従うと仮定する：

（v）$c(\mathbf{x} + \mathbf{y}) = c\mathbf{x} + c\mathbf{y}$　（ベクトルの分配律）

（vi）$(c_1 + c_2)\mathbf{x} = c_1\mathbf{x} + c_2\mathbf{x}$　（スカラーの分配律）

（vii）$c_1(c_2\mathbf{x}) = (c_1c_2)\mathbf{x}$　（結合律）

（viii）$e\mathbf{x} = \mathbf{x}.$　（e はスカラーの体の単位要素とする）

公理（i）から（viii）までを満足する（スカラーの体 F と結合した）ベクトルの集合を，**線形ベクトル空間** $\mathscr{V}$，より正確には，$\mathscr{V}(F)$ または $\mathscr{V}_F$ とよぶ．条件（iii）と（iv）を1つの条件にまとめると，任意の2つの要素 $\mathbf{x}$ と $\mathbf{y}$ にたいして，$\mathbf{x} + \mathbf{z} = \mathbf{y}$ を満足する要素 $\mathbf{z}$ が一意的に存在することになる．

ある与えられた k にたいして，順序づけられた実数の組 $(x_1, \cdots, x_k)$ の全体が (1a.1.1) と (1a.1.2) で定義した加法とスカラー倍をもつなら，それはベクトル空間を構成することを読者は容易に確かめうるであろう．次数が k 以下のすべての多項式についても，係数は任意の体に属し，加法と定数との積がふつうのように定義される場合には，同じことが成り立つ．この本で述べる統計的方法の取り扱いにおいては，順序づけられた実数の組としてのベクトルを主として考えるけれども，その公理論的な表現は，新しい代数についての正しい見通しを読者に与え，かつ，高度な統計的方法の研究に次第に用いられるようになってきたヒルベルト空間 (Halmos (1951) を見よ) のような，もっと複雑なベクトル空間の研究の基礎を準備することになる．

ベクトル空間 $\mathscr{V}$ の**線形部分空間**，**部分空間**，または，**線形多様体**とは，加法とスカラー倍に関して閉じた，任意の部分集合 $\mathscr{M}$ のことである．閉じたというのは，$\mathbf{x}$ と $\mathbf{y}$ が $\mathscr{M}$

に属するとき，スカラーの任意の対 c, d にたいして，$(c\mathbf{x}+d\mathbf{y}) \in \mathscr{M}$ であることをいう．そのような部分集合 $\mathscr{M}$ は，それ自身，$\mathscr{V}$ におけると同じ定義の加法とスカラー倍に関する1つのベクトル空間である．零ベクトルしか含まない部分集合，および，$\mathscr{V}$ のすべての要素から成る部分集合は，部分空間の両極端の例である．この2つは特異な部分空間とよばれ，それら以外は正則な部分空間である．

一例として，ベクトル $\alpha_1, \cdots, \alpha_k$ よりなる一定の集合 S のあらゆる1次結合をとると，これは，S によって張られた線形多様体 $\mathscr{M}(S)$ と呼ばれる部分空間になる．これが S を含む最小の部分空間である．

n 変数 $x_1, \cdots, x_n$ の k 個の1次方程式，

$$a_{i1}x_1 + \cdots + a_{in}x_n = 0, \qquad i = 1, \cdots, k$$

を考える．ここで，a_{ij} は任意の体 F に属するとする．この解 $(x_1, \cdots, x_n)$ をベクトルと考えると，その全体は (1a. 1. 1) と (1a. 1. 2) で定義した加法とスカラー積をもつ部分空間をつくることを，読者は確かめることができるであろう．

1a. 2　ベクトル空間の基底

一組のベクトル $\mathbf{u}_1, \cdots, \mathbf{u}_k$ は，$c_1\mathbf{u}_1 + \cdots + c_k\mathbf{u}_k = 0$ を満足する，すべてが同時にはゼロでないスカラー $c_1, \cdots, c_k$ が存在するとき，1次従属であるといい，そうでないとき，1次独立であるという．そのように定義すると，次が成り立つ：

1. 零ベクトル自身は，1つの従属な組である．
2. 零ベクトルを含むベクトルの組は，いずれも従属である．
3. 零でないベクトル $\mathbf{u}_1, \cdots, \mathbf{u}_k$ の組は，その1つのベクトルがそれ以外のものの1次結合であるときにかぎって，従属となる．

ベクトル空間 $\mathscr{V}$ において，$\mathscr{V}$ を生成する，すなわち $\mathscr{V}$ を張る，ベクトルの1次独立な部分集合を，$\mathscr{V}$ の基底（ハーメル基底）とよぶ．

（i）すべてのベクトル空間 $\mathscr{V}$ は基底をもつ．

これを証明するために，$\mathscr{V}$ の零でないベクトル $\alpha_1, \alpha_2, \cdots$ を次々に選び，どの α_i もそれより前に選ばれたものに従属しないようにする．この過程で，その k 番目のあとには $\mathscr{V}$ に独立なベクトルが1つも残らないようなことが起こるかもしれない．この場合には，$\alpha_1, \cdots, \alpha_k$ が基底をつくり，$\mathscr{V}$ は有限 k 次元のベクトル空間であるといわれる．他方，α_i を選ぶ過程にかぎりがない場合には，$\mathscr{V}$ は無限次元をもつといわれる．あらゆるベクトルを生成するには，無限個の独立なベクトルの組の存在が必要であることを示すためには，さらに論証が必要である．しかし，われわれの研究分野は有限次元のベクトル空間にかぎられるから，この証明は省略する．有限次元の空間に関する，次の諸結果は重要である．

（ii）$\alpha_1, \cdots, \alpha_k$ と $\beta_1, \cdots, \beta_s$ がともに基底として選ばれたものなら，$s = k$ でなければならない．

いま，$s > k$ としよう．組 $\beta_1, \alpha_1, \cdots, \alpha_k$ を考えると，これは従属である．もし α_i がその他のものに従属であるなら，$\mathscr{V}$ は，$\beta_1, \alpha_1, \cdots, \alpha_{i-1}, \alpha_{i+1}, \cdots, \alpha_k$ によっても生成されることになり，その場合には，組 $\beta_2, \beta_1, \alpha_1, \cdots, \alpha_{i-1}, \alpha_{i+1}, \cdots, \alpha_k$ が従属となる．そこで，もう1つの α を除くことができる．このように1つの β を加え，1つの α を除くという操作を，組 $\beta_1, \cdots, \beta_k$ が残されるまで続けることができる（どの段階でも β は除けないことに注意せよ）．そして，この場合には，$\beta_1, \cdots, \beta_k$ が $\mathscr{V}$ を張り，したがって，残りの $s-k$ 個の β ベクトルは余分となる．すべての基底に共通な基数 k は，その空間を張るのに必要なベクトルの最小数，または，その空間における独立なベクトルの最大数をあらわすことになる．この数を，$\mathscr{V}$ の次元 dimension または位数 rank とよび，$d[\mathscr{V}]$ であらわす．

（iii）$\mathscr{V}$ のベクトルはすべて，所与の基底によって一意的に表現される．

もし同じベクトルが，$\sum a_i\alpha_i$ と $\sum b_i\alpha_i$ によってあらわされるならば，$\sum(a_i - b_i)\alpha_i = 0$ となり，α_i は独立であるから，この関係はすべての i について $a_i - b_i = 0$ でなければ成立しない．

$x_i \in F$（要素の1つの体）の k 個の要素 $(x_1, \cdots, x_k)$ の順序づけられたすべての組よりなるユークリッド空間 $E_k(F)$ は特に興味がある．そのベクトルは，k 次元ユークリッド空間の点と考えられる．k 個のベクトル $\mathbf{e}_1 = (1, 0, \cdots, 0)$，$\mathbf{e}_2 = (0, 1, 0, \cdots, 0)$，$\cdots$，$\mathbf{e}_k = (0, 0, \cdots, 1)$ は，E_k のなかにあり，独立で，任意のベクトルは，$\mathbf{x} = (x_1, \cdots, x_k) = x_1\mathbf{e}_1 + \cdots + x_k\mathbf{e}_k$ と書ける．それゆえ，$d[E_k] = k$ で，ベクトル $\mathbf{e}_1, \cdots, \mathbf{e}_k$ は基底をつくり，したがって，独立な k 個のベクトルの他のどんな組も基底をつくる．それゆえ，E_k の任意のベクトルは，k 個の独立なベクトルの1次結合として一意的にあらわされ，そして，当然のことながら，k 個の独立なベクトルを含む任意のベクトルの組の（必ずしも一意的ではない）結合としてもあらわされる．

F が実数体であるときは，ベクトル空間は単に R^k とあらわされる．実験計画の研究では，有限個の要素をもつガロア体 (GF) が用いられるから，そのベクトル空間には有限個のベクトルしか含まれない．そのような空間にたいしては，記号 $E_k(GF)$ が用いられる．F が複素数体のときのベクトル空間 $E_k(F)$ は，U^k であらわされ，k 次元ユニタリ空間とよばれる．**1a. 3** 節の取り扱いは，任意の F にたいしても有効である．のちに，われわれの考察の対象を R^k だけにしぼるであろう．(iv) では，有限次元ベクトル空間の研究は E_k の研究と等価であることを示す重要な結果を証明する．

（iv）$d[\mathscr{V}] = k$ の任意のベクトル空間 $\mathscr{V}_F$ は，E_k と同型である．

$\alpha_1, \cdots, \alpha_k$ を $\mathscr{V}_F$ の基底とすると，要素 $\mathbf{u} \in \mathscr{V}_F$ は，$a_i \in F$，$i = 1, \cdots, k$ を用いて，

$\mathbf{u} = a_1\alpha_1 + \cdots + a_k\alpha_k$ とあらわされる．すると，対応

$$\mathbf{u} \to \mathbf{u}^* = (a_1, \cdots, a_k), \qquad \mathbf{u}^* \in E_k(F)$$

によって，同型が確立される．なぜなら，もし $\mathbf{u} \to \mathbf{u}^*$, $\mathbf{v} \to \mathbf{v}^*$ なら，$\mathbf{u} + \mathbf{v} \to \mathbf{u}^* + \mathbf{v}^*$, $c\mathbf{u} \to c\mathbf{u}^*$ となるからである．この結果は，同じ次元をもつ任意の2つのベクトル空間は同型であることを示している．

1a.3　1次方程式

$\alpha_1, \cdots, \alpha_m$ を任意のベクトル空間 $\mathscr{V}_F$ の m 個の定数ベクトルとし，スカラー $x_1, \cdots, x_m \in F$（付与される要素の体）についての1次方程式

$$x_1\alpha_1 + \cdots + x_m\alpha_m = 0 \tag{1a.3.1}$$

を考える．

（i）(1a.3.1) が，自明ではない解，すなわちすべての x_i が同時に零ではないような解をもつための必要十分条件は，$\alpha_1, \cdots, \alpha_m$ が従属であることである．

（ii）この解を $E_m(F)$ における（行）ベクトル $\mathbf{x} = (x_1, \cdots, x_m)$ と考えると，それは1つのベクトル空間をつくる．

もし $\mathbf{x}$ と $\mathbf{y}$ がともに解なら，$a\mathbf{x} + b\mathbf{y}$ も解となるから，これは正しい．α_i 自身がベクトル空間 $\mathscr{V}_F$ に属していることに注意せよ．

（iii）$\mathscr{S}$ を解の線形部分空間，$\mathscr{M}$ をベクトル $\alpha_1, \cdots, \alpha_m$ の線形部分空間とせよ．すると，$d[\mathscr{S}] = m - d[\mathscr{M}]$ となる．ここで，記号 d はその空間の次元を示す．

一般性を失うことなく，$\alpha_1, \cdots, \alpha_k$ は独立である，すなわち，$d[\mathscr{M}] = k$ とすると，その場合には

$$\alpha_j = a_{j1}\alpha_1 + \cdots + a_{jk}\alpha_k, \qquad j = k+1, \cdots, m$$

と書ける．このとき，ベクトル

$$\begin{aligned} \beta_1 &= (a_{k+1,1}, \cdots, a_{k+1,k}, -1, 0, \cdots, 0) \\ &\cdots\cdots\cdots\cdots \\ \beta_{m-k} &= (a_{m,1}, \cdots, a_{m,k}, 0, 0, \cdots, -1) \end{aligned} \tag{1a.3.2}$$

は独立で，いずれも (1a.3.1) を満足することがわかる．(1a.3.2) で与えられる組が解のすべてを張るならば，これは解空間 $\mathscr{S}$ の基底となるであろう．いま，$\mathbf{y} = (y_1, \cdots, y_m)$ を任意の解とし，ベクトル

$$\mathbf{y} + y_{k+1}\beta_1 + \cdots + y_m\beta_{m-k} \tag{1a.3.3}$$

を考えると，これもまた解となる．しかし，これは，$(z_1, \cdots, z_k, 0, \cdots, 0)$ の形の解であって，$z_1\alpha_1 + \cdots + z_k\alpha_k = 0$ が満足される．この関係は，$z_1 = \cdots = z_k = 0$ でなければ成立しない．したがって，(1a.3.3) は零ベクトルとなり，任意の解 $\mathbf{y}$ は，$m-k$ 個の独立なベクトル $\beta_1, \cdots, \beta_{m-k}$ であらわされることになる．ここで，k は独立な α ベクトルの数である．よって上の結果は証明された．

（iv）非同次方程式

$$x_1\alpha_1 + \cdots + x_m\alpha_m = \alpha_0 \tag{1a.3.4}$$

は，α_0 が $\alpha_1, \cdots, \alpha_k$ に従属するとき，そしてそのときのみ，解をもつ．

この結果は，ベクトルの従属性の定義から出る．

（v）非同次方程式 (1a.3.4) の一般解は，(1a.3.4) の任意の特殊解と，同次方程式 (1a.3.1) の一般解の和である．

この結果は，(1a.3.4) の任意の2つの解の差は，(1a.3.1) の解であることから出る．

(1a.3.4) の解の組は疑いもなく E_m に属するが，同次方程式 (1a.3.1) の解の場合とは違って，部分空間をつくらないことに注意せよ．(v) から，次の結果 (vi) が導かれる．

（vi）非同次方程式の解は，対応する同次方程式が零ベクトルを唯一の解とするとき，すなわち $d[\mathscr{S}] = m - k = 0$，いいかえると $\alpha_1, \cdots, \alpha_m$ が独立のとき，一意的に定まる．

1次方程式の特殊で重要な場合は，$\alpha_i \in E_n$ のときである．いま，$\alpha_i = (a_{1i}, \cdots, a_{ni})$ とおくと，(1a.3.1) は次のように書ける：

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1m}x_m &= 0 \\ &\cdots\cdots\cdots\cdots \\ a_{n1}x_1 + \cdots + a_{nm}x_m &= 0, \end{aligned} \tag{1a.3.5}$$

これは，m 個の未知数についての n 個の1次方程式である．式 (1a.3.5) の列ベクトルが $\alpha_1, \cdots, \alpha_m$ である．この行ベクトルを $\beta_i = (a_{i1}, \cdots, a_{im})$ であらわす．いま，r 行だけが独立であるとし，列ベクトルによって生成される部分空間の次元を $k = d[\mathscr{M}(\alpha)]$ とする．もし r 個の独立な行だけを残しても，得られる解空間 $\mathscr{S}$ は，同じものになるであろう．$d[\mathscr{S}] = m - k$ であるから，方程式の残された組では，どれか k 個の列が独立となる．ところで，列ベクトルはいまや r 個の要素しかもたない．すなわち，E_r に属するから，独立な列ベクトルの数は r を超えることはできない．それゆえ $r \geqq k$ となる．この議論で列と行の役割を取り替えることができるから，$k \geqq r$ も得られる．したがって，$r = k$ となり，次が得られる：

（vii）(1a.3.5) で係数 a_{ij} だけを考えると，独立な行の数は，独立な列の数に等しく，

$$d[\mathscr{M}(\alpha)] = d[\mathscr{M}(\beta)], \qquad d[\mathscr{S}] = m - d[\mathscr{M}(\alpha)] = m - d[\mathscr{M}(\beta)]$$

である．

式 (1a.3.5) の右辺の $(0, \cdots, 0)$ を $\alpha_0 = (a_1, \cdots, a_n)$ におきかえて得られる非同次方程式を考えよう．解の存在を証明するためには，次の等価な必要かつ十分条件のどれか1つが正しいことを確かめる必要がある．

1. (iv) により，列ベクトル α_0 は $\alpha_1, \cdots, \alpha_m$ に従属である．すなわち，

$$d[\mathscr{M}(\alpha_1, \cdots, \alpha_m)] = d[\mathscr{M}(\alpha_1, \cdots, \alpha_m, \alpha_0)].$$

2. β_i に第 $m+1$ 番目の要素として a_i を加えて得られる $m+1$ 次元ベクトルを γ_i とする．このとき，$d[\mathscr{M}(\gamma_1, \cdots, \gamma_n)] = d[\mathscr{M}(\beta_1, \cdots, \beta_n)]$ となる．この条件は，その前者が列についての同じ値 $d[\mathscr{M}(\alpha_1, \cdots, \alpha_m, \alpha_0)]$ に等しく，後者が $d[\mathscr{M}(\alpha_1, \cdots, \alpha_m)]$ に等しいことから明らかである．

3. $c_1, \cdots, c_n$ が $c_1\beta_1 + \cdots + c_n\beta_n = 0$ を満足するとき，$c_1a_1 + \cdots + c_na_n = 0$ となる，あるいは，同じことであるが，$c_1\gamma_1 + \cdots + c_n\gamma_n = 0$ となる．これを証明するには，$\sum c_i\gamma_i = 0$ のどの解も明らかに $\sum c_i\beta_i = 0$ の解であり $d[\mathscr{M}(\beta)] = d[\mathscr{M}(\gamma)]$ であるという結果からわかるように，$\sum c_i\beta_i = 0$ と $\sum c_i\gamma_i = 0$ の解 $(c_1, \cdots, c_n)$ はまったく同じであることを見さえすればよい．

4. 行ベクトル $\beta_1, \cdots, \beta_n$ が独立であれば，非同次方程式は任意の α_0 にたいして1つの解をもつ．

1a. 4　内積をもつベクトル空間

これまでは，ベクトル空間の要素のあいだの関係を，独立性の概念を通して研究してきた．2次元や3次元のユークリッド幾何学の研究に出てくる，ベクトル間の距離や角度というような別の概念を導入することも，これからは役に立つであろう．付与されるスカラーが実数あるいは複素数であるときには，これらの概念はすぐベクトル空間に拡張され，次の諸条件を満足する，2つのベクトル $\mathbf{x}, \mathbf{y}$ の内積 $(\mathbf{x}, \mathbf{y})$ として知られる複素数関数によってあらわされる：

(a)　$(\mathbf{x}, \mathbf{y}) = \overline{(\mathbf{y}, \mathbf{x})}$，　$\overline{(\mathbf{y}, \mathbf{x})}$ は $(\mathbf{x}, \mathbf{y})$ の共役複素数

(b)　$\mathbf{x} \neq 0$ なら　$(\mathbf{x}, \mathbf{x}) > 0$，　$\mathbf{x} = 0$ なら　$= 0$

(c)　$(t\mathbf{x}, \mathbf{y}) = t(\mathbf{x}, \mathbf{y})$，$t$ はスカラー

(d)　$(\mathbf{x} + \mathbf{y}, \mathbf{z}) = (\mathbf{x}, \mathbf{z}) + (\mathbf{y}, \mathbf{z})$.

実数座標をもつ n 次元ユークリッド空間 R^n にたいしては，実数関数 $(\mathbf{x}, \mathbf{y}) = \sum x_iy_i$ が条件 (a)〜(d) を満足する．この関数を R^n におけるベクトルの内積ときめる．複素数座標をもつユニタリ空間 U^n にたいしては，$(\mathbf{x}, \mathbf{y})$ を $\sum x_i\bar{y}_i$ と定義する．実際，U^n におけるあらゆる内積は，a_{ij} を適当に選ぶと，$\sum\sum a_{ij}x_i\bar{y}_j$ の形に書けることが示される（本章末の問題 19. 5 参照）．

$(\mathbf{x}, \mathbf{x})$ の正の平方根を $\|\mathbf{x}\|$ であらわし，$\mathbf{x}$ のノルム，または $\mathbf{x}$ の長さと呼ぶ．関数 $\|\mathbf{x} - \mathbf{y}\|$ は，測度空間における距離のもつ公理を満足するから，$\mathbf{x}$ と $\mathbf{y}$ との距離と考えられる．三角不等式

$$\|\mathbf{x} - \mathbf{y}\| + \|\mathbf{y} - \mathbf{z}\| \geqslant \|\mathbf{x} - \mathbf{z}\|$$

および，コーシー＝シュヴァルツの不等式

$$|(\mathbf{x}, \mathbf{y})| \leqslant \|\mathbf{x}\|\|\mathbf{y}\|$$

を証明すると興味深いであろう．

2つのベクトル $\mathbf{x}, \mathbf{y}$ は，その内積が $(\mathbf{x}, \mathbf{y}) = 0$ のとき，直交するといわれる．零でないベクトル $\mathbf{x}, \mathbf{y} \in R^k$ のあいだの角度 θ は

$$\cos\theta = \frac{(\mathbf{x}, \mathbf{y})}{\|\mathbf{x}\| \cdot \|\mathbf{y}\|}$$

によって定義され，この $\cos\theta$ は閉区間 $[-1, 1]$ のなかに存在する．以下に，内積が定義されたベクトル空間のもつ諸性質を調べよう．

(i)　一組の零でないベクトル $\alpha_1, \cdots, \alpha_k$ は，2つずつ互いに直交するとき，必然的に独立となる．

もしそうでなければ，$\sum a_i\alpha_i = 0$ で，$a_j \neq 0$ とせよ．α_j と $\sum a_i\alpha_i$ との内積をとり，$i \neq j$ のとき $(\alpha_i, \alpha_j) = 0$ とおくと，$a_j(\alpha_j, \alpha_j) = 0$ を得，これから $(\alpha_j, \alpha_j) > 0$ であるから $a_j = 0$ が導かれて，仮定と矛盾する．

(ii)　グラム＝シュミットの直交化　ベクトル空間 $\mathscr{V}$ を張る直交ベクトルの組を直交基底 o. b. (orthogonal basis) と呼ぶ．さらに，各ベクトルのノルムが1のときは，正規直交基底 o. n. b. (orthonormal basis) とよばれる．o. n. b. はつねに存在すると言える．$\alpha_1, \alpha_2, \cdots$ を任意の基底とし，次々に，

$$\begin{aligned} \xi_1 &= \alpha_1 \\ \xi_2 &= \alpha_2 - a_{21}\xi_1 \\ &\cdots\cdots\cdots\cdots \\ \xi_i &= \alpha_i - a_{i,i-1}\xi_{i-1} - \cdots - a_{i1}\xi_1 \\ &\cdots\cdots\cdots\cdots \end{aligned} \tag{1a. 4. 1}$$

をつくれ．ただし，i 番目の式の係数は，ξ_i と，$\xi_1, \cdots, \xi_{i-1}$ の各々との内積がゼロになるようにきめられる．こうして $\xi_1, \cdots, \xi_{i-1}$ をきめたあとは，ξ_i を定義する係数 $a_{i1}, \cdots, a_{i,i-1}$ は次のようにして求められる：

$$(\alpha_i, \xi_1) = a_{i1}(\xi_1, \xi_1), \quad \cdots, \quad (\alpha_i, \xi_{i-1}) = a_{i,i-1}(\xi_{i-1}, \xi_{i-1})$$

このベクトル $\xi_1, \xi_2, \cdots$ は o. b. である．$\eta_i = \xi_i/\|\xi_i\|$，$i = 1, 2, \cdots$ で定義される組が o. n. b. になる．ここで $\|\xi_i\| \neq 0$ である．なぜなら，もしそうでなければ，$\xi_i = 0$ となり，$\alpha_1, \cdots, \alpha_i$ のあいだに従属性があることになる．また，$\xi_1, \cdots, \xi_i$ は，$\alpha_1, \cdots, \alpha_i$ で張られる線形部分空間の o. b. となり，$\mathscr{M}(\xi_1, \cdots, \xi_i) = \mathscr{M}(\alpha_1, \cdots, \alpha_i)$，$i = 1, 2, \cdots$ となる．o. n. b. の存在を証明するこの構成方法は，グラム＝シュミットの直交化法として知られている．

(iii)　ベクトルの直交射影　$\mathscr{W}$ を有限次元部分空間，α を $\mathscr{W}$ に属さないベクトルと

する．このとき，2つのベクトル γ, β が存在して，$\alpha=\beta+\gamma$ と書け，$\beta\in\mathscr{W}$, $\gamma\neq 0$ は $\mathscr{W}$ に直交する（$\gamma\perp^{\mathrm{er}}\mathscr{W}$）ようにできる．さらに，この γ と β は一意的に定まる．

$\alpha_1,\cdots,\alpha_r$ を $\mathscr{W}$ の基底とする．このとき，$\alpha_1,\cdots,\alpha_r,\alpha_{r+1}$ $(=\alpha)$ から，グラム＝シュミットの直交化法によって，$\xi_1,\cdots,\xi_r,\xi_{r+1}$ をつくる．すると，(1a. 4. 1) の最後の式から，

$$\alpha=\xi_{r+1}+(a_{r+1,r}\xi_r+\cdots+a_{r+1,1}\xi_1)=\xi_{r+1}+\beta=\gamma+\beta$$

が得られ，これによって γ と β を定義する．$\mathscr{W}=\mathscr{M}(\alpha_1,\cdots,\alpha_r)=\mathscr{M}(\xi_1,\cdots,\xi_r)$ であるから，$\beta\in\mathscr{W}$ で，$\xi_{r+1}=\gamma$ は，その構成法から見て，$\xi_1,\cdots,\xi_r$ の各々と直交する．したがって，部分空間 $\mathscr{M}(\xi_1,\cdots,\xi_r)$ すなわち $\mathscr{W}$ に直交することがわかる．一方，$\beta_1+\gamma_1$ という別の分解があるとしよう．このときは $0=(\gamma-\gamma_1)+(\beta-\beta_1)$ となる．両辺と $\gamma-\gamma_1$ との内積をとると，$(\gamma-\gamma_1,\gamma-\gamma_1)=0$ を得，$\gamma-\gamma_1=0$ であることがわかり，したがって，$\beta-\beta_1=0$ にもなる．ベクトル γ は α から $\mathscr{W}$ への $\perp^{\mathrm{er}}$（垂線），β は α の $\mathscr{W}$ 上への射影とよばれる．$\alpha\in\mathscr{W}$ なら $\gamma=0$ となる．

（iv）α から $\mathscr{W}$ への $\perp^{\mathrm{er}}$ である γ は次の性質をもつ：

$$\|\gamma\|=\inf\|\alpha-\mathbf{x}\|,\qquad \mathbf{x}\in\mathscr{W} \tag{1a. 4. 2}$$

$\alpha=\beta+\gamma$, $\beta\in\mathscr{W}$, $\gamma\perp^{\mathrm{er}}\mathscr{W}$ とする．このとき，$\mathbf{x}\in\mathscr{W}$ にたいして，$(\mathbf{x}-\beta,\gamma)=0$. したがって，

$$\begin{aligned}(\mathbf{x}-\alpha,\mathbf{x}-\alpha)&=(\mathbf{x}-\beta-\gamma,\mathbf{x}-\beta-\gamma)\\&=(\mathbf{x}-\beta,\mathbf{x}-\beta)+(\gamma,\gamma)\geqslant\|\gamma\|^2\\&=\|\gamma\|^2\qquad(\mathbf{x}=\beta\text{ のとき})\end{aligned}$$

が証明できる．それゆえ，α と $\mathscr{W}$ の要素の間の最短距離は α から $\mathscr{W}$ への $\perp^{\mathrm{er}}$ の長さである．

（v）ベッセルの不等式　$\alpha_1,\cdots,\alpha_m$ を，内積をもつあるベクトル空間の正規直交系，β を他の任意のベクトルとする．このとき，

$$|(\alpha_1,\beta)|^2+\cdots+|(\alpha_m,\beta)|^2\leqslant\|\beta\|^2 \tag{1a. 4. 3}$$

が成り立ち，等号は，β が $\alpha_1,\cdots,\alpha_m$ によって生成される部分空間に属するときにのみ成立する．

$\delta=\sum(\alpha_i,\beta)\alpha_i$ とおく．すると，δ は β の $\mathscr{M}(\alpha_1,\cdots,\alpha_m)$ 上への射影となり，命題（iii）によって，$\beta=\delta+\gamma$ と書け，$(\delta,\gamma)=0$ となる．したがって，

$$\|\beta\|^2=\|\delta\|^2+\|\gamma\|^2\geqslant\|\delta\|^2=\sum|(\alpha_i,\beta)|^2$$

（vi）所与のベクトルの組 S は部分空間であってもなくてもよいが，これに直交するすべてのベクトルの組は部分空間を構成し，$S^\perp$ であらわされる．

この結果は，部分空間の定義から得られる．ベクトル空間全体 $\mathscr{V}$ は n（有限）次元であり，S によって張られる部分空間 $\mathscr{M}(S)$ の基底を $\alpha_1,\cdots,\alpha_r$ とする．いま，$S^\perp$ の基底を $\beta_1,\cdots,\beta_s$ とすると，どの β_i もすべての α_j に $\perp^{\mathrm{er}}$ であるから，ベクトルの組 $\alpha_1,\cdots,\alpha_r,\beta_1,\cdots,\beta_s$ は独立である．明らかに，$(r+s)\not>n$. いま，$(r+s)<n$ とすると，$\alpha_1,\cdots,\alpha_r,\beta_1,\cdots,\beta_s$ に独立なベクトルが少なくとも1つ存在することになるから，（iii）によってそれらのすべてに直交する，零でないベクトルをみつけることができる．そのベクトルは，$\alpha_1,\cdots,\alpha_r$ の各々と必ず直交するから，$S^\perp$ に属さねばならず，したがって，$\beta_1,\cdots,\beta_s$ に従属することになり，矛盾を生じる．それゆえ，$r+s=n$ である．こうして，次の重要な結果が得られる．

$$d[\mathscr{M}(S)]+d[S^\perp]=d[\mathscr{V}] \tag{1a. 4. 4}$$

さらに，$\mathscr{V}$ のベクトルはどれも，$\mathbf{x}+\mathbf{y}$, $\mathbf{x}\in\mathscr{M}(S)$, $\mathbf{y}\in S^\perp$ の形に書くことができる．

（vii）**部分空間の直和**　ベクトル空間 $\mathscr{V}$ が，部分空間 $\mathscr{V}_1,\cdots,\mathscr{V}_k$ の直和であると言われ，$\mathscr{V}=\mathscr{V}_1\oplus\cdots\oplus\mathscr{V}_k$ と書けるのは，$\alpha\in\mathscr{V}$ が，$\alpha_i\in\mathscr{V}_i$ を用いて，一意的に，$\alpha_1+\cdots+\alpha_k$ とあらわされる場合である．そのような場合には，

$$\mathscr{V}_i\cap\mathscr{V}_j=\{0\}\quad\text{かつ}\quad d(\mathscr{V})=d(\mathscr{V}_1)+\cdots+d(\mathscr{V}_k) \tag{1a. 4. 5}$$

が成立する．

（vi）で示した分解 $\mathscr{V}=\mathscr{M}(S)+S^\perp$ は直和の一例である．

補足と問題

1　$\alpha_1,\cdots,\alpha_k$ をベクトル空間 $\mathscr{V}$ における k 個の独立なベクトルとするとき，それらを広げてつねに $\mathscr{V}$ の基底がつくれる，すなわち $r=d[\mathscr{V}]$ のとき，$\alpha_{k+1},\cdots,\alpha_r$ を加えて，$\alpha_1,\cdots,\alpha_r$ が $\mathscr{V}$ の基底をつくるようにできる，ことを示せ．

2　次の行ベクトルによって生成される空間の次元は4である．

1	1	0	1	0	0
1	1	0	0	1	0
1	1	0	0	0	1
1	0	1	1	0	0
1	0	1	0	1	0
1	0	1	0	0	1

［ヒント：行および列によって形成される空間の次元は同じである．明らかに独立であるさいごの4列を考え，はじめの2列はこれらに従属することに注目せよ．］

3　問題2の結果を，pq 個の行ベクトルが，それぞれ p 行よりなる q ブロックにまとめられている場合に一般化せよ．ここで，第 i ブロックの第 j ベクトルは，1番目，$i+1$ 番目，および $q+j+1$ 番目の列で1で，残りの $p+q-2$ 列では0とする．このように生成された空間の次元は $p+q-1$ であることを示せ．

4　pq 個の要素 a_{ij}, $i=1,\cdots,p$; $j=1,\cdots,q$ は，すべての i,j,r,s にたいして，その4つの差 $a_{ij}+a_{rs}-a_{is}-a_{rj}=0$ であるならば，$a_1,\cdots,a_p$ と $b_1,\cdots,b_q$ を適当に選ぶと，a_{ij} は $a_{ij}=a_i+b_j$ という形に表現できる．

［問題2,3,4は，第4章，4e節の分散分析による2元分類データの研究に役立つ．］

5　$(a,b,\cdots,b)$, $(b,a,b,\cdots,b)$, $\cdots$, $(b,b,\cdots,a)$ という組で，各ベクトルの要素の和が

ゼロであるとき，独立なベクトルの数を求めよ．

6　$\mathscr{S}$ と $\mathscr{T}$ は部分空間とし，$\mathscr{S}$ と $\mathscr{T}$ に共通なベクトルの集合を，$\mathscr{S}$ と $\mathscr{T}$ との交わり $\mathscr{S}\cap\mathscr{T}$ で，$\mathbf{x}\in\mathscr{S}$，$\mathbf{y}\in\mathscr{T}$ なるベクトル $\mathbf{x}+\mathbf{y}$ の集合を $\mathscr{S}+\mathscr{T}$ であらわす．このとき，(i) $\mathscr{S}\cap\mathscr{T}$ と $\mathscr{S}+\mathscr{T}$ もまた部分空間である，(ii) $\mathscr{S}\cap\mathscr{T}$ は $\mathscr{S}$ と $\mathscr{T}$ の両方に含まれる最大の部分空間である，(iii) $\mathscr{S}+\mathscr{T}$ は，$\mathscr{S}$ と $\mathscr{T}$ の双方を含む最小の部分空間である，(iv) $d[\mathscr{S}+\mathscr{T}]=d[\mathscr{S}]+d[\mathscr{T}]-d[\mathscr{S}\cap\mathscr{T}]$ である，ことを示せ．

7　R^n は $r\leqslant n$ なる r 個の部分空間の直和であらわされることを示せ．

8　S をベクトルの任意の集合，$S^\perp$ を S に直交するすべてのベクトルの集合とする．$(S^\perp)^\perp\supset S$ で，S が部分空間であるときのみ，$(S^\perp)^\perp=S$ であることを示せ．

9　S と T は，ベクトルの任意の2つの集合で，零ベクトルを含むとする．このとき，$(S+T)^\perp=S^\perp\cap T^\perp$ が成り立つ．

10　$\mathscr{S}$ と $\mathscr{T}$ を部分空間とすると，$(\mathscr{S}\cap\mathscr{T})^\perp=\mathscr{S}^\perp+\mathscr{T}^\perp$ である．

11　$\alpha_1,\cdots,\alpha_m$ を任意のベクトル空間 $\mathscr{V}_F$ のベクトルとする．いま，$(t_1,\cdots,t_m)$ が E_m のすべてのベクトルの上を動くとき，ベクトル $t_1\alpha_1+\cdots+t_m\alpha_m$，$t_i\in F$ の集合は，線形多様体 $\mathscr{M}(\alpha)$ をつくり，その次元は，組 $\alpha_1,\cdots,\alpha_m$ のなかの独立なベクトルの数 k に等しくなる．ところで，$(t_1,\cdots,t_m)$ の動きが E_m のなかの s 次元の部分空間 $\mathscr{S}$ に制限されるときには，集合 $t_1\alpha_1+\cdots+t_m\alpha_m$ について何を言うことができるか．制限された集合もまた部分空間 $\subset\mathscr{M}(\alpha)$ であり，この空間の次元は $s-d[\mathscr{S}\cap\mathscr{T}]$ となることを示せ．ただし，$\mathscr{T}$ は，方程式 $t_1\alpha_1+\cdots+t_m\alpha_m=0$ の解 $(t_1,\cdots,t_m)$ をあらわす E_m の部分空間である．

12　問題**11**の α_i は E_n に属するとせよ．すなわち，

$$\alpha_i=(a_{i1},\cdots,a_{in}),\qquad i=1,\cdots,m$$

と書け，その列ベクトルを

$$\beta_j=(a_{1j},\cdots,a_{mj}),\qquad j=1,\cdots,n$$

であらわす．$\mathbf{t}=(t_1,\cdots,t_m)$ は，E_m での固定したベクトル $\gamma_1,\cdots,\gamma_k$ にたいして，式

$$(\mathbf{t},\gamma_i)=0,\qquad i=1,\cdots,k$$

を満足するとする．すると，$\sum t_i\alpha_i$ によって生成される部分空間の次元は，

$$d[\mathscr{M}(\beta_1,\cdots,\beta_n,\gamma_1,\cdots,\gamma_k)]-d[\mathscr{M}(\gamma_1,\cdots,\gamma_k)]$$

となることを示せ．

13　次の方程式

$$\begin{aligned}2x_1-x_2+5x_3&=4,\\4x_1\qquad+6x_3&=1,\\-2x_2+4x_3&=7+\delta\end{aligned}$$

が解をもつための δ の値を求めよ．

14　次の方程式，

$$\begin{aligned}4x_1+3x_2+2x_3+8x_4&=\delta_1\\3x_1+4x_2+x_3+7x_4&=\delta_2\\x_1+x_2+x_3+x_4&=\delta_3\end{aligned}$$

は，任意の $\delta_1,\delta_2,\delta_3$ にたいして解をもつ．［3つの方程式があって，3つの行はすべて独立であることに注意せよ．］この解の全体を定め，そのうちで最小の長さをもつものを求めよ．

15　0, 1, 2は，剰余類 (mod 3) をあらわすとする．それらは，3つの要素から成るガロア体，GF (3) をつくり，剰余類の加法と乗法の通常の定義が適用される．この系のなかで，

$$\begin{aligned}x_1+x_2+x_3&=0\\x_1+2x_2\qquad&=0\end{aligned}$$

の解の基底を求め，かつ，方程式の行ベクトルに依存する解が存在するか否かを調べよ．［そのようなことは，実数の方程式を考えているときには可能ではないことに注意せよ．］また，次の方程式を解け．

$$\begin{aligned}x_1+2x_2+0+x_4&=1\\2x_1+0+x_3+2x_4&=2\end{aligned}$$

16　$\alpha_1,\cdots,\alpha_m$ によって張られる部分空間への $\mathbf{x}$ の射影は，$a_1\alpha_1+\cdots+a_m\alpha_m$ によって与えられることを示せ．ただし，$a_1,\cdots,a_m$ は，方程式

$$\begin{aligned}a_1(\alpha_1,\alpha_1)+\cdots+a_m(\alpha_1,\alpha_m)&=(\mathbf{x},\alpha_1)\\\cdots\cdots\cdots\cdots\\a_1(\alpha_m,\alpha_1)+\cdots+a_m(\alpha_m,\alpha_m)&=(\mathbf{x},\alpha_m)\end{aligned}$$

の任意の解である．これは，$(\mathbf{x}-a_1\alpha_1-\cdots-a_m\alpha_m)$ が，$\alpha_1,\cdots,\alpha_m$ の各々と $\perp^{\text{er}}$ であるための条件を表示することによって得られる．この方程式は，射影が存在する以上，解をもつ（別な方法としては，**1a.3**に与えた方程式が解をもつための条件を用いて証明せよ）．さらに，$\sum a_i\alpha_i$ は一意的に定まり，また，解 $(a_1,\cdots,a_m)$ が何であろうとも，垂線の長さ $a_1(\mathbf{x},\alpha_1)+\cdots+a_m(\mathbf{x},\alpha_m)$ も一意的に定まる．

17　$\alpha_1,\cdots,\alpha_m$ を R^k のベクトルとする．$p_1,\cdots,p_m$ を所与のスカラーとするとき，条件 $(\mathbf{b},\alpha_1)=p_1,\cdots,(\mathbf{b},\alpha_m)=p_m$ の下で，$(\mathbf{b},\mathbf{b})=\sum b_i^2$ の最小値を求めよ．

18　スタイニッツの取替問題　$\alpha_1,\cdots,\alpha_k$ をベクトル空間 $\mathscr{V}$ の基底，$\beta_1,\cdots,\beta_p$ を $\mathscr{V}$ の $p\ (<k)$ 個の独立なベクトルとするとき，α ベクトルのなかから $k-p$ 個のベクトル $\alpha_{(1)},\cdots,\alpha_{(k-p)}$ を選んで，$\beta_1,\cdots,\beta_p,\alpha_{(1)},\cdots,\alpha_{(k-p)}$ が $\mathscr{V}$ の基底となるようにすることができる．

19　α と β を正規化したベクトルとする．このとき，$|(\alpha,\beta)|=1\Rightarrow\alpha=c\beta$，ただし $|c|=1$．

1b　行列と行列式の理論

1b.1　行　列　演　算

行列とは，pq 個の要素（任意の体に属する）を p 行 q 列に配列したものをいう．各要素は，それが現れる行と列の番号をあらわす対 (i,j) によって見分けられる．行列は，一般に，$\mathbf{A}$ のような大文字であらわされ，必要に応じて，より明白に，記号 (a_{ij}) によってあらわされる．ここに，a_{ij} は (i,j) 要素である．行列には次の演算が定義される．

加法　$\mathbf{A}=(a_{ij})$ と $\mathbf{B}=(b_{ij})$ を，同じ次数 $p\times q$，すなわち，それぞれ p 行 q 列

をもつ2つの行列とすると，**A** と **B** の和は，同じ次数 $p \times q$ をもつ行列 $\mathbf{A}+\mathbf{B}=(a_{ij}+b_{ij})$ であると定義される．次が容易に証明できる：

$$\mathbf{A}+\mathbf{B}=\mathbf{B}+\mathbf{A}$$

$$\mathbf{A}+(\mathbf{B}+\mathbf{C})=(\mathbf{A}+\mathbf{B})+\mathbf{C}=\mathbf{A}+\mathbf{B}+\mathbf{C}$$

スカラー倍　ベクトルについてと同様に，行列にもスカラー c をかけることができ，その結果は，同じ形の行列

$$c\mathbf{A}=(ca_{ij})$$

すなわち，各要素にスカラー c をかけたものになる．

$$(c+d)\mathbf{A}=c\mathbf{A}+d\mathbf{A}, \qquad c(\mathbf{A}+\mathbf{B})=c\mathbf{A}+c\mathbf{B}$$

となることを確かめよ．

行列のかけ算　積 **AB** は，**A** の列の数が **B** の行の数に等しいときにだけ定義される．その結果は，その (i,j) 要素が，**A** の i 番目の行ベクトルと **B** の j 番目の列ベクトルの積である行列となる．$\mathbf{A}=(a_{ij})$，$\mathbf{B}=(b_{ij})$ とすると，その積は公式

$$\mathbf{AB}=(c_{ij}), \qquad c_{ij}=\sum_r a_{ir}b_{rj}$$

によって求められる．**A** が $p \times q$，**B** が $q \times t$ であるなら，$\mathbf{C}=\mathbf{AB}$ は $p \times t$ となる．積行列の次数を定める公式，$(p \times q)(q \times t)=(p \times t)$ に注意せよ．次を確かめよ：

$$\begin{pmatrix} -1 & 2 & 3 \\ 5 & 1 & 2 \end{pmatrix}\begin{pmatrix} 1 & 2 & 0 \\ 4 & 1 & 3 \\ 3 & 1 & 2 \end{pmatrix}=\begin{pmatrix} 16 & 3 & 12 \\ 15 & 13 & 7 \end{pmatrix}.$$

<u>行列のかけ算には，交換律が成立しない．</u>すなわち，<u>一般に，$\mathbf{AB} \neq \mathbf{BA}$ である．</u>**AB** は定義できるが，**BA** は定義できないということも起こる．しかしながら，結合律は満足される．すなわち，

$$\mathbf{A}(\mathbf{BC})=(\mathbf{AB})\mathbf{C}=\mathbf{ABC}$$

行列 **A, B, C** のかけ算ができるためには，それらの次数は $p \times q$，$q \times r$，$r \times s$ でなければならず，その場合には **ABC** の次数は $p \times s$ となる．分配律もまた成立する．

$$\mathbf{A}(\mathbf{B}+\mathbf{C})=\mathbf{AB}+\mathbf{AC}$$

$$(\mathbf{A}+\mathbf{B})(\mathbf{C}+\mathbf{D})=\mathbf{A}(\mathbf{C}+\mathbf{D})+\mathbf{B}(\mathbf{C}+\mathbf{D}).$$

零行列と単位行列　**0** であらわされる零行列は，その要素がすべてゼロの行列である．次数が $q \times q$ の正方行列で，その対角線要素がすべて1，非対角線要素が全部0のものは，単位行列といわれる．そのような行列を **I**，または，その次数を示すことが必要なときには，$\mathbf{I}_q$ であらわす．

$$\mathbf{A}+\mathbf{0}=\mathbf{A}$$

および，積が定義できるときには，

$$\mathbf{IA}=\mathbf{A}, \qquad \mathbf{AI}=\mathbf{A}$$

となることを確かめよ．そうならば，行列 **0** と **I** は，数体系における要素0と1と同じ役割を果たす．

行列の転置　行列 **A** の行と列を交換して得られる行列 $\mathbf{A}'$（$\mathbf{A}^T$ であらわすこともある）を **A** の転置行列という．これは，$\mathbf{A}'$ の (i,j) 要素が **A** の (j,i) 要素に等しいことを意味する．定義から，次を得る．

$$(\mathbf{AB})'=\mathbf{B}'\mathbf{A}', \qquad (\mathbf{ABC})'=\mathbf{C}'\mathbf{B}'\mathbf{A}', \qquad \text{など}.$$

R^n のベクトルを列方向に書くと，次数 $n \times 1$ の行列となる．いま，**X** と **Y** を2つの列ベクトルとすると，その内積 $(\mathbf{X},\mathbf{Y})$ は，行列記号では，$\mathbf{X}'\mathbf{Y}$ と書くことができる．**A** を $m \times n$ 行列とすると，**AX** は，**A** の列ベクトルの1次結合である1つの列ベクトルになることを注意せよ．これから先では，列ベクトルを単にベクトルと呼ぶことにしよう．

共役転置　これは，**A** の要素が複素数体に属するという特殊の場合に定義される．行列 $\mathbf{A}^*$ の (i,j) 要素が行列 **A** の (j,i) 要素の共役複素数であるとき，$\mathbf{A}^*$ を **A** の共役転置という．定義から，

$$(\mathbf{AB})^*=\mathbf{B}^*\mathbf{A}^*, \qquad (\mathbf{ABC})^*=\mathbf{C}^*\mathbf{B}^*\mathbf{A}^*, \qquad \text{など}.$$

行列の階数　行列は，特定の順序で書かれた行ベクトル（または列ベクトル）の一組であると考えられる．すると，行列の階数 rank とは，独立な行の数である，あるいは，同じ数になるが，独立な列の数である，と定義される［(vii), **1a.3** を見よ］．

行列の逆　次数 $m \times m$ の正方行列は，その階数が m のとき，正則であるという．そのような場合には，**A** の逆行列として周知の $\mathbf{A}^{-1}$ が一意的に存在し，

$$\mathbf{AA}^{-1}=\mathbf{A}^{-1}\mathbf{A}=\mathbf{I} \tag{1b.1.1}$$

が成り立つ．これを証明するために，$\mathbf{A}^{-1}=(b_{ij})$ が存在するとする．$\mathbf{A}=(a_{ij})$ と書くと，$(a_{ij})(b_{ij})=\mathbf{I}$．この左辺を展開し，両辺の第1列の要素を等しいとおくと，$b_{11}, b_{21}, \cdots, b_{m1}$（$\mathbf{A}^{-1}$ の第1列の要素）についての連立1次方程式が得られる．

$$\begin{aligned} a_{11}b_{11}+\cdots+a_{1m}b_{m1}&=1 \\ a_{21}b_{11}+\cdots+a_{2m}b_{m1}&=0 \\ \cdots\cdots\cdots\cdots\cdots\cdots & \\ a_{m1}b_{11}+\cdots+a_{mm}b_{m1}&=0 \end{aligned} \tag{1b.1.2}$$

(a_{ij}) は最大階数 full rank をもつから，式 (1b.1.2) の解は一意的である［(vi), **1a.3**］．$\mathbf{A}^{-1}$ の他の列も同様に，$\mathbf{AA}^{-1}=\mathbf{I}$ になるように定められる．まったく同じ論法で，$\mathbf{CA}=\mathbf{I}$ なるような行列 **C** をみつけることができる．この後から $\mathbf{A}^{-1}$ をかけると，

$$\mathbf{CAA}^{-1}=\mathbf{IA}^{-1}=\mathbf{A}^{-1} \Rightarrow \mathbf{CI}=\mathbf{A}^{-1} \Rightarrow \mathbf{C}=\mathbf{A}^{-1}$$

を得，(1b.1.1) が証明されたことになる．**A** と **B** を $\mathbf{A}^{-1}$ と $\mathbf{B}^{-1}$ が存在するような行列と

すると，$(\mathbf{AB})^{-1}=\mathbf{B}^{-1}\mathbf{A}^{-1}$ となる．これは，$\mathbf{ABB}^{-1}\mathbf{A}^{-1}=\mathbf{I}$ であることからわかる．どの行列も逆が存在するときは，

$$(\mathbf{ABC})^{-1}=\mathbf{C}^{-1}\mathbf{B}^{-1}\mathbf{A}^{-1}$$

などが成立することを確かめよ．また，$(\mathbf{A}')^{-1}=(\mathbf{A}^{-1})'$，および $(\mathbf{A}^*)^{-1}=(\mathbf{A}^{-1})^*$ であることがわかる．

次数 $m\times n$ の行列 $\mathbf{B}$ を考えよう．$\mathbf{BC}=\mathbf{I}_m$ なる行列 $\mathbf{C}$ が存在するとき，$\mathbf{C}$ を $\mathbf{B}$ の右正則逆行列とよび，$\mathbf{B}_R^{-1}$ であらわす．同様に，左正則逆行列 $\mathbf{B}_L^{-1}$ が存在することもある．

右逆行列は〔$\mathbf{B}$ の階数〕$R(\mathbf{B})=m\leqslant n$ のときのみ存在する．その場合には，(1b. 1. 2) のように $\mathbf{C}$ の列要素についての方程式を書くことができる．解の存在は [(iv), 1a. 3] よりわかる．同様に，左逆行列は，$R(\mathbf{B})=n\leqslant m$ のときのみ存在する．

ユニタリ行列　要素が複素数の正方行列 $\mathbf{A}$ は，$\mathbf{A}^*\mathbf{A}=\mathbf{I}$ が成り立つとき，ユニタリと呼ばれる．その場合には，$\mathbf{A}^*=\mathbf{A}^{-1}$ で，$\mathbf{AA}^*=\mathbf{I}$ となる．$\mathbf{A}$ の要素が実数のときは，$\mathbf{A}'\mathbf{A}=\mathbf{AA}'=\mathbf{I}$ なら，$\mathbf{A}$ は直交行列とよばれる．

分割行列　1つの行列を，2つ以上の行列に分割して並べるのが便利なことがある．このとき，分割行列は

$$\mathbf{A}=\left(\begin{array}{c:c}\mathbf{P}&\mathbf{Q}\\ \hdashline \mathbf{R}&\mathbf{S}\end{array}\right)=\begin{pmatrix}\mathbf{P}&\mathbf{Q}\\ \mathbf{R}&\mathbf{S}\end{pmatrix}\qquad(1b.\ 1.\ 3)$$

とあらわされ，誤解の恐れがなければ，点線はなくてもよい．ただし，(1b. 1. 3) では，$\mathbf{P}$ の行数は $\mathbf{Q}$ の行数に等しく，$\mathbf{P}$ の列数は $\mathbf{R}$ の列数に等しいことなどが成立していなければならない．転置の定義から，

$$\mathbf{A}'=\begin{pmatrix}\mathbf{P}'&\mathbf{R}'\\ \mathbf{Q}'&\mathbf{S}'\end{pmatrix}.$$

2つの分割行列の積は，行列を要素として扱い，これに行と列とのかけ算の規則を適用することによって得られる．したがって，たとえば，

$$\mathbf{AB}=\begin{pmatrix}\mathbf{P}&\mathbf{Q}\\ \mathbf{R}&\mathbf{S}\end{pmatrix}\begin{pmatrix}\mathbf{E}\\ \mathbf{G}\end{pmatrix}=\begin{pmatrix}\mathbf{PE}+\mathbf{QG}\\ \mathbf{RE}+\mathbf{SG}\end{pmatrix}$$

$$\mathbf{AC}=\begin{pmatrix}\mathbf{P}&\mathbf{Q}\\ \mathbf{R}&\mathbf{S}\end{pmatrix}\begin{pmatrix}\mathbf{E}&\mathbf{F}\\ \mathbf{G}&\mathbf{H}\end{pmatrix}=\begin{pmatrix}\mathbf{PE}+\mathbf{QG}&\mathbf{PF}+\mathbf{QH}\\ \mathbf{RE}+\mathbf{SG}&\mathbf{RF}+\mathbf{SH}\end{pmatrix}$$

もちろん，$\mathbf{PE}$ などの積が存在する場合についてである．

正方行列 $\mathbf{A}$ は，$\mathbf{A}=\mathbf{A}'$ のとき対称行列，$\mathbf{A}=\mathbf{A}^*$ のときエルミート行列と呼ばれる．

1b. 2　基本行列と行列の対角化

次数 $m\times n$ の行列 $\mathbf{A}$ について，次の演算を考えよう．

(e_1) = 1つの行（または1つの列）にスカラー c をかけること．

(e_2) = r 番目の行（または列）を，r 番目の行（列）と s 番目の行（列）の λ 倍との和におきかえること．

(e_3) = 2つの行（または列）を交換すること．

これらの演算によって行列の階数が変らないことは明らかである．これらの演算は正則な行列をかけるのと同じであることを以下に示そう．r 番目の行にスカラー c をかけるという演算 (e_1) は，クロネッカー Δ として周知の次数 $m\times m$ の正方行列を，$\mathbf{A}$ の前からかけることに等しい．ここで，クロネッカー Δ とは，対角要素にのみ0とは異なる値をとり得るが，他はすべて零の正方行列であって，ここでは，その要素が特に次の値をとる場合にあたる．

$$\Delta=(\delta_{ij}),\qquad \delta_{rr}=c;\quad \delta_{ii}=1,\quad i\neq r;\quad \delta_{ij}=0,\quad i\neq j.$$

行についての演算 (e_2) は，次に定義する次数 $m\times m$ の正方行列 $\mathbf{E}_{rs}(\lambda)$ を $\mathbf{A}$ の前からかけることと同じである．

$\mathbf{E}_{rs}(\lambda)=(e_{ij})$,　$e_{ii}=1$ すべての i について；$e_{rs}=\lambda$；それ以外は $e_{ij}=0$.

列についての同様の演算にたいしては，Δ および $\mathbf{E}_{sr}(\lambda)$ 型の $n\times n$ 正方行列を後からかける必要がある．r 番目と s 番目の2行を交換する演算 (e_3) は，$\mathbf{A}$ の前から行列 $\mathbf{E}_{sr}(-1)$, $\mathbf{E}_{rs}(1)$, $\mathbf{E}_{sr}(-1)$, Δ を次々にかけるのに等しいことを確かめよ．同様に，列の交換は，後から次々に行列をかけることに等しい．行列のより単純な形への標準化に関する次の諸結果は重要である．

(i)　**台形**　<u>$\mathbf{A}$ を次数 $m\times n$ の行列とする．$\mathbf{BA}$ が台形になるような $m\times m$ 正則行列 $\mathbf{B}$ が存在する．</u>

台形の行列とは，(a)　各行は，全部の要素が0であるか，または，0から始って最初に1が現れ，その後は必ずしも0でなくてもよい，かつ，(b)　行の最初の1を含む列はどれも他の要素は0である，ときそしてこの場合のみをいう．

第1行の最初の零でない要素（もしあれば）を選び，これが第 i 列にあるとする．第1行を，第1枢軸要素とよばれるこの選ばれた要素で割る．こうして得られた行に定数をかけて第2行から引き，第2行の i 番目の列の要素を零にすることができる．同様の演算によって，i 番目の列の要素は枢軸の位置にあるもの以外はすべて零にする（すなわち，掃き出す）ことができる．次に第2行に移り，同じ演算を繰り返す．ある段階で，行が零ばかりから成るときは，その次の行に進む．こうして得られる行列が台形になること，しかも，それが (e_1), (e_2) 型の演算だけで得られることが容易にわかる．したがって，この標準化は，一連の Δ または $\mathbf{E}_{rs}(\lambda)$ 型の正方行列を前からかけることによって達成され，その積は正確に $\mathbf{B}$ になる．Δ, $\mathbf{E}_{rs}(\lambda)$ 型の行列は，これらをかけても階数は変わらず，かつ，各行列の階数は m であるから，その積も同じ階数 m をもつ．それゆえ，$\mathbf{B}$ は正則であ

る．

（ii）エルミート正準形　Aを次数mの正方行列とする．このとき，$\mathbf{CA}=\mathbf{H}$ とするような次数mの正則行列 $\mathbf{C}$ が存在する．ただし，$\mathbf{H}$ は，$\mathbf{H}^2=\mathbf{H}$（ベキ等）という性質を満足するエルミート正準形である．

エルミート正準形とは，行列の主対角要素が0と1ばかりから成りその対角要素の下の要素はすべて0のものをいう．すなわち，対角要素が0のときはその行の要素は全部0であり，対角要素が1のときはその列のほかの要素は全部0であるようなときに得られる．

$\mathbf{BA}$ を台形にするような行列を，$\mathbf{B}$ とする．この $\mathbf{BA}$ を $\mathbf{H}$ 型に標準化するには，いくつかの行を交換するだけでよい．しかも，行の交換は，(e_3) 型の演算，すなわち，基本行列を前からかけることによってできる．$\mathbf{BA}$ を $\mathbf{H}$ 型に標準化するのに用いる基本行列の積を $\mathbf{G}$ とすると，$\mathbf{A}$ を $\mathbf{H}$ 型に標準化するのに必要な行列は $\mathbf{C}=\mathbf{GB}$ と書ける．明らかに，$R(\mathbf{A})=R(\mathbf{CA})=R(\mathbf{H})$ で，この数は $\mathbf{H}$ の対角要素中の1の数に等しい［$\mathbf{A}$ の階数に $R(\mathbf{A})$ という記号を用いた］．$\mathbf{H}^2=\mathbf{H}$ となることは容易に確かめられる．

エルミート正準化は，1次方程式

$$\mathbf{AX}=\mathbf{Y} \tag{1b.2.1}$$

を解くのに有用であって，特に $\mathbf{A}$ が特異のときにそうである．(1b.2.1) の両辺に $\mathbf{C}$ をかけると，それと同等な方程式 $\mathbf{HX}=\mathbf{CY}$ を得，この方程式が解をもつのは，$\mathbf{H}$ の i 番目の対角要素が0のとき $\mathbf{CY}$ の i 番目の要素が0である場合にかぎられる．そのような場合の解は $\mathbf{X}=\mathbf{CY}$ となる．$\mathbf{A}$ の階数が最大でないとき，行列 $\mathbf{C}$ は $\mathbf{A}$ の一般逆行列と呼ばれ，$\mathbf{A}^-$ であらわされる．一般逆行列についてのより詳細な議論は **1b.5** で行う．

（iii）対角化　Aを階数 r の $m\times n$ 行列とする．このとき，次数がそれぞれmとnの正則な正方行列 $\mathbf{B}$ と $\mathbf{C}$ が存在して，

$$\mathbf{BAC}=\begin{pmatrix}\mathbf{I}_r & 0\\ 0 & 0\end{pmatrix} \tag{1b.2.2}$$

とできる．ここで，$\mathbf{I}_r$ は次数 r の単位行列で，それ以外の要素はすべて0である．

$\mathbf{A}$ の台形標準化から出発し，任意の行の1の要素を枢軸として，その後に0でない要素があれば，そのすべてを掃き出す．これは，列についての (e_2) 型の演算，すなわち，1つの列の定数倍を別の列に加えること，つまり，基本行列を後からかけることによって達成できる．この結果得られた形で，行の交換と列の交換をしさえすれば，主対角要素に1をもってくることができる．行列 $\mathbf{C}$ は後からのかけ算に用いられたすべての基本行列の積であり，$\mathbf{B}$ は前からのかけ算に用いられた行列の積となる．

（iv）階数分解　(1b.2.2) からまた，階数 r の $m\times n$ 行列 $\mathbf{A}$ は次の形に書ける．すなわち，

$$\begin{aligned}\mathbf{A}&=\mathbf{B}^{-1}\begin{pmatrix}\mathbf{I}_r & 0\\ 0 & 0\end{pmatrix}\mathbf{C}^{-1}\\ &=\mathbf{B}_1\mathbf{C}_1\end{aligned} \tag{1b.2.3}$$

ここで，$\mathbf{B}_1, \mathbf{C}_1$ は $m\times r$ および $r\times n$ の行列で，いずれも階数 r をもつ．

$\mathbf{A}$ を階数 r の $m\times n$ 行列とする．このとき，$m\times m$ の正則行列 $\mathbf{M}$ と $n\times n$ の直交行列 $\mathbf{N}$ が存在して，

$$\mathbf{A}=\mathbf{M}\begin{pmatrix}\mathbf{I}_r & 0\\ 0 & 0\end{pmatrix}\mathbf{N}$$

と書ける．

この結果は容易に証明できる．

（v）三角化　Aを次数mの正方行列とする．このとき，$m\times m$ の正則行列 $\mathbf{B}$ が存在して，$\mathbf{BA}$ を三角形，すなわち，対角線の下の要素はすべて0，にできる．

$\mathbf{A}$ の第1対角要素は0でないとすると，この第1行に定数をかけ，他の行から差し引くことによって，第1列を掃き出すことができる．第1対角要素が0ならば，第1要素が0でない任意の行を第1行に加え，それから列を掃き出せばよい．第1列の要素が全部0ならば，第2行に進め．第1行と第1列を無視して，同じ演算を繰り返せ．すべての演算は (e_2) 型のものだけである．前からかけた行列の積が $\mathbf{B}$ になる．

（vi）対称行列の対角化　Aを対称な正方行列とする．このとき，正則行列 $\mathbf{B}$ が存在して，$\mathbf{BAB}'$ を対角形にすることができる．すなわち，$\mathbf{C}$ を正則，Δ を対角行列とすると $\mathbf{A}$ は積 $\mathbf{C}\Delta\mathbf{C}'$ としてあらわされる．

$\mathbf{A}$ の第1対角要素を考える．それが0でなければ，それを枢軸として第1行と第1列の要素を掃き出した行列を得ることができる．この掃き出しは，基本行列を前からと後からと次々にかけることに等しく，対称性のために，前後の行列は互いに転置したものになる．こうして得られた行列もまた対称である．この標準化した行列の第1対角要素を考え，それが0でなければ，これを用いて，第2行第2列にある0でない要素を掃き出す．この操作は，すでに標準化した第1行第1列を変えない．どの段階における標準化した行列でも，その第1対角要素が0でなければ，この操作をつづけることができる．もし対応する行と列の要素がすべて0であれば，その次の対角要素を考え，前と同様に進む．あるいは，適当な行とそれに対応する列のスカラー倍を加える（または，引く）ことによって，標準化した行列の第1の位置に零でない要素をもってき，前述の操作をつづけることができる．これもまた，基本行列を用いた，行と列について対称な演算である．前からのかけ算に用いた基本行列の積として $\mathbf{B}$ を選ぶことにより，上の結果が得られる．最後に，$\mathbf{C}=\mathbf{B}^{-1}$ である．

（vii）　直交行列による三角化　　$\mathbf{A}$ は $m \times n$ 行列で，$m \geqq n$ とする．このとき，次数 $m \times m$ の直交行列 $\mathbf{B}$ が存在し，$\mathbf{BA} = \begin{pmatrix} \mathbf{T} \\ 0 \end{pmatrix}$ とできる．ここで，$\mathbf{T}$ は次数 n の上側三角行列，0 は次数 $(m-n) \times n$ の，0 ばかりから成る行列である．

$\mathbf{u}$ を長さ 1 の任意のベクトルとすると，$(\mathbf{I} - 2\mathbf{uu}')$ は直交行列であることがわかる．ここで，行列 $(\mathbf{I} - 2\mathbf{uu}')\mathbf{A}$ の第 1 列は，その第 1 要素をあるいは例外とするほかは，すべての要素が 0 となるように，$\mathbf{u}$ を定めることができることを示そう．このためには，方程式

$$(\mathbf{I} - 2\mathbf{u}_1\mathbf{u}_1')\alpha_1 = \beta$$

を解く必要がある．ただし，α_1 は $\mathbf{A}$ の第 1 列，β はその第 1 要素 β_1 があるいは 0 ではないかもしれないが，残りの要素はすべて 0 のベクトルである．$\alpha_1 = 0$，あるいは，その第 1 要素だけが 0 でない場合には，$\mathbf{u}_1 = 0$ とすればよいことはすぐにわかる．そうでないときには，

$$\left.\begin{aligned} \beta_1 &= \sqrt{\alpha_1'\alpha_1} \\ \mathbf{u}_1 &= \frac{\alpha_1 - \beta}{2b} \end{aligned}\right\} \qquad (1\text{b}.2.4)$$

を計算せよ．ただし，$4b^2 = (\alpha_1 - \beta)'(\alpha_1 - \beta)$ である．(1b. 2. 4) のように $\mathbf{u}_1$ を選ぶと，$\mathbf{A}$ に $(\mathbf{I} - 2\mathbf{u}_1\mathbf{u}_1')$ をかけた結果は，その第 1 列が，第 1 要素以外は 0 になる（掃き出される）ことを確かめよ．この第 1 列第 1 行を除いて標準化した行列を考える．同様の演算をこの行列の第 1 列に，次数 $(m-1) \times (m-1)$ の直交行列 $(\mathbf{I} - 2\mathbf{u}_2\mathbf{u}_2')$ を用いて施こす．これは，$(\mathbf{I} - 2\mathbf{u}_1\mathbf{u}_1')\mathbf{A}$ の左から，同じく直交行列である分割行列

$$\begin{pmatrix} 1 & 0' \\ 0 & \mathbf{I} - 2\mathbf{u}_2\mathbf{u}_2' \end{pmatrix}$$

をかけることに等しい．この演算を，すべての列が掃き出されるまで繰り返すと，$\begin{pmatrix} \mathbf{T} \\ 0 \end{pmatrix}$ 型の行列に到達する．このとき，定理で述べられた行列 $\mathbf{B}$ は，直交行列の積となり，それ自身も直交になる．この標準化の方法は Householder (1964) による．

（viii）　QR 分解　　$\mathbf{A}$ は階数 r の $m \times n$ 実行列で，$m \geqq n$ とする．このとき，$\mathbf{Q}'\mathbf{Q} = \mathbf{I}$ という部分直交な $m \times n$ 行列 $\mathbf{Q}$ と，上側三角形の $n \times n$ 行列 $\mathbf{R}$ が存在して，

$$\mathbf{A} = \mathbf{QR} \qquad (1\text{b}.2.5)$$

と分解できる．

この結果は，(vii) の行列 $\mathbf{B}$ の最初の n 行によって形成される部分行列として $\mathbf{Q}'$ を選び，$\mathbf{T}$ を $\mathbf{R}$ と書くと得られる．

$R(\mathbf{A}) = n$ で，$\mathbf{R}$ の対角要素はすべて非負（つねに可能）であることが要求されるなら，この分解は一意的であることに注意せよ．$R(\mathbf{A}) = n$ のときに 2 つの分解があると，$+1$ と -1 より成る対角行列 $\mathbf{S}$ を用いて，$\mathbf{Q}_1 = \mathbf{Q}_2\mathbf{S}$，$\mathbf{R}_1 = \mathbf{SR}_2$ と関係づけられる．

（ix）　グラム＝シュミットの三角化　　$\mathbf{A}$ を $m \times n$ 行列とする．このとき，$n \times n$ の上側三角行列 $\mathbf{T}$ と，その列が正規直交系の $m \times n$ 行列 $\mathbf{S}$ が存在して，$\mathbf{A} = \mathbf{ST}$ と書ける．

$\alpha_1, \cdots, \alpha_n$ を $\mathbf{A}$ の列，$\sigma_1, \cdots, \sigma_n$ を $\mathbf{S}$ の列，$t_{1i}, \cdots, t_{ii}$ を $\mathbf{T}$ の i 番目の列の対角線の上側の要素とする．このとき，[(ii)，1a. 4] で述べたグラム＝シュミットの直交化法におけるように，

$$\alpha_i = t_{1i}\sigma_1 + \cdots + t_{ii}\sigma_i$$

で，$j = 1, \cdots, i-1$ にたいして $t_{ji} = \sigma_j'\alpha_i$，かつ

$$t_{ii} = (\alpha_i'\alpha_i - t_{1i}^2 - \cdots - t_{i-1,i}^2)^{1/2}$$

とできる．それゆえ，係数 t_{ij} とベクトル $\sigma_1, \sigma_2, \cdots$ は次のように，逐次に求められる：

$$\alpha_1 = t_{11}\sigma_1$$
$$\alpha_2 = t_{12}\sigma_1 + t_{22}\sigma_2$$
$$\cdots\cdots\cdots\cdots$$

ここで，$t_{ii} = 0$ のときは，$j > i$ なる j にたいして $t_{ij} = 0$ とし，σ_i は任意に選ぶと約束する．

1b. 3　行　列　式

次数 m の正方行列 $\mathbf{A} = (a_{ij})$ の行列式は，要素 a_{ij} の実数値関数で，

$$|\mathbf{A}| = \sum \pm a_{1i}a_{2j}\cdots a_{mp}$$

によって定義される．ここで，和は，$(1, 2, \cdots, m)$ のあらゆる順列 $(i, j, \cdots, p)$ についてとられ，$(i, j, \cdots, p)$ が偶順列のときプラス，奇順列のときマイナスの符号をとる．行列の行列式の公理的な導出については，Levi (1942) と Rao (1952) を見よ．

要素 a_{ij} の余因子を記号 A_{ij} であらわし，第 i 行と第 j 列を除いて得られる行列式の $(-1)^{i+j}$ 倍であると定義する．次の性質を証明するのは容易である．

（a）　$R(\mathbf{A}) \neq m$ のとき，またそのときのみ，$|\mathbf{A}| = 0$ となる．

（b）　任意の 2 行（または 2 列）を交換すると，$|\mathbf{A}|$ の符号は変わる．

（c）　$|\mathbf{A}| = \sum_i a_{ri}A_{ri}$　　任意の r について
　　$= \sum_r a_{ri}A_{ri}$　　任意の i について．

（d）　$r \neq s$ のとき，　$\sum_i a_{ri}A_{si} = 0$.

（e）　$|\mathbf{A}| \neq 0$ ならば，$\mathbf{A}^{-1} = (A_{ij}/|\mathbf{A}|)'$.

（f）　$\mathbf{A}$ が対角行列または三角行列ならば，$|\mathbf{A}|$ は対角要素の積となる．

（g）　$\mathbf{A}$ と $\mathbf{B}$ が正方行列なら，$|\mathbf{AB}| = |\mathbf{A}||\mathbf{B}|$.

1b.4 線 形 変 換

著者によっては，行列理論を線形変換を通して取り扱っている．E_n（n 次元ユークリッド空間）から E_m への線形変換は，$\boldsymbol{\xi}, \mathbf{X} \in E_n$，$\boldsymbol{\eta}, \mathbf{Y} \in E_m$，$a, b \in F$，$F$ はベクトル空間 E_n と E_m とに付与されたスカラーの体，とすると，$\mathbf{X} \to \mathbf{Y}$ でかつ $\boldsymbol{\xi} \to \boldsymbol{\eta}$ なら，$a\mathbf{X} + b\boldsymbol{\xi} \to a\mathbf{Y} + b\boldsymbol{\eta}$ に変換されるという性質をもっている．以下に，線形変換と行列のあいだの対応を示そう．

基本ベクトル $\mathbf{e}_1, \cdots, \mathbf{e}_n \in E_n$ を次のように変換する（E_n から E_m への）線形変換を考える：

$$\mathbf{e}_1' = (1, 0, \cdots, 0) \to (a_{11}, \cdots, a_{m1}),$$

$$\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\mathbf{e}_n' = (0, 0, \cdots, 1) \to (a_{1n}, \cdots, a_{mn}).$$

$\mathbf{X}' = (x_1, \cdots, x_n) = x_1\mathbf{e}_1' + \cdots + x_n\mathbf{e}_n' \in E_n$ とする．線形性により

$$\mathbf{X}' \to (\sum a_{1i}x_i, \cdots, \sum a_{mi}x_i) = \mathbf{Y}'$$

となるから，これは，$\mathbf{X} \to \mathbf{Y}$ が，$\mathbf{A}$ を $m \times n$ 行列

$$\mathbf{A} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ & \cdots\cdots\cdots & \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

とするとき，$\mathbf{Y} = \mathbf{AX}$ という関係によって得られることと同等である．

$\mathbf{Z} = \mathbf{BY}$ で，かつ $\mathbf{Y} = \mathbf{AX}$ なら，$\mathbf{Z} = (\mathbf{BA})\mathbf{X}$ は，2つの行列の積を含む，$\mathbf{X}$ から $\mathbf{Z}$ への直接の変換を与える．

$\mathbf{X}$ から $\mathbf{Y}$ への変換は1対1である必要はない．$\mathbf{A}$ が正則で $\mathbf{A}^{-1}$ が存在するときだけその変換は1対1となり，その逆変換は $\mathbf{X} = \mathbf{A}^{-1}\mathbf{Y}$ で与えられる（1対1の対応を示す）．

直交変換　線形変換 $\mathbf{Y} = \mathbf{CX}$ は，$\mathbf{CC}' = \mathbf{I}$ のとき，直交と言われる．$|\mathbf{C}|^2 = 1$ となるから，この変換は正則である．条件 $\mathbf{CC}' = \mathbf{I}$ から，$\mathbf{C}$ の行ベクトルは互いに直交し，その長さは1である．いいかえれば，それらは R^n 空間の o. n. b. を構成することがわかる．

$$|\mathbf{C}| = \pm 1, \quad \mathbf{C}' = \mathbf{C}^{-1}, \quad \text{それゆえ，} \mathbf{C}'\mathbf{C} = \mathbf{I}$$

を証明することは容易である．この最後の結果から，列もまた o. n. b. をつくることがわかる．

（i）<u>直交変換の重要な性質は，距離と角度を不変に保つことである．</u>

$\mathbf{X} \to \mathbf{Y}$，$\boldsymbol{\xi} \to \boldsymbol{\eta}$ とすると，

$$(\mathbf{Y} - \boldsymbol{\eta})'(\mathbf{Y} - \boldsymbol{\eta}) = (\mathbf{CX} - \mathbf{C}\boldsymbol{\xi})'(\mathbf{CX} - \mathbf{C}\boldsymbol{\xi})$$
$$= (\mathbf{X} - \boldsymbol{\xi})'\mathbf{C}'\mathbf{C}(\mathbf{X} - \boldsymbol{\xi}) = (\mathbf{X} - \boldsymbol{\xi})'(\mathbf{X} - \boldsymbol{\xi})$$

となるから，2点間の距離は不変である．同様に，

$$(\mathbf{Y}_1 - \boldsymbol{\eta}_1)'(\mathbf{Y}_2 - \boldsymbol{\eta}_2) = (\mathbf{X}_1 - \boldsymbol{\xi}_1)'(\mathbf{X}_2 - \boldsymbol{\xi}_2)$$

となるから，ベクトル間の角度もまた不変である．実際，距離が不変であるための必要十分条件は，変換が直交であることである．

射影　R^n における線形変換の一例は，部分空間 $\mathscr{M}$ へのベクトルの射影である．なぜなら，$\mathbf{Y}_1, \mathbf{Y}_2$ を $\mathbf{X}_1, \mathbf{X}_2$ の射影とすると，$a\mathbf{Y}_1 + b\mathbf{Y}_2$ は $a\mathbf{X}_1 + b\mathbf{X}_2$ の射影となるからである．したがって，射影は，行列 $\mathbf{P}$，つまり，$\mathbf{Y} = \mathbf{PX}$ を通じて獲得される．$\mathbf{P}$ の明白な表現は，**1c.4** で与える．

1b.5 一 般 逆 行 列

定義　任意の階数をもつ $m \times n$ 行列 $\mathbf{A}$ を考える．$\mathbf{A}$ の一般逆行列（g-inverse）は，$n \times m$ 行列で，$\mathbf{A}^-$ であらわすと，任意の $\mathbf{Y} \in \mathscr{M}(\mathbf{A})$ にたいして，$\mathbf{X} = \mathbf{A}^-\mathbf{Y}$ が方程式 $\mathbf{AX} = \mathbf{Y}$ の解となるものをいう．

一般逆行列の存在を証明し，その性質をしらべよう．

（i）<u>$\mathbf{A}^-$ が一般逆行列である $\Leftrightarrow \mathbf{AA}^-\mathbf{A} = \mathbf{A}$.</u>

$\mathbf{Y}$ として $\mathbf{A}$ の i 番目の列 $\mathbf{a}_i$ を選ぶ．方程式 $\mathbf{AX} = \mathbf{a}_i$ は明らかに解をもち，$\mathbf{X} = \mathbf{A}^-\mathbf{a}_i$ が解となる．よって，$\mathbf{AA}^-\mathbf{a}_i = \mathbf{a}_i$ がすべての i について成り立ち，これは $\mathbf{AA}^-\mathbf{A} = \mathbf{A}$ と同じになる．したがって，$\mathbf{A}^-$ の存在 $\Rightarrow \mathbf{AA}^-\mathbf{A} = \mathbf{A}$ がいえた．逆に，$\mathbf{A}^-$ が $\mathbf{AA}^-\mathbf{A} = \mathbf{A}$ を満足し，かつ，$\mathbf{AX} = \mathbf{Y}$ が解をもつなら，$\mathbf{AA}^-\mathbf{AX} = \mathbf{AX}$，すなわち $\mathbf{AA}^-\mathbf{Y} = \mathbf{Y}$ を得る．これは，$\mathbf{X} = \mathbf{A}^-\mathbf{Y}$ が解であることである．それゆえ，一般逆行列の別の定義として，

$$\mathbf{AA}^-\mathbf{A} = \mathbf{A} \tag{1b.5.1}$$

を満足する任意の行列 $\mathbf{A}^-$ ということができる．

(1b.5.1) で定義した一般逆行列は，必ずしも一意的ではない．そこで，著者によっては，(1b.5.1) 以外にさらにいくつかの条件をつけた行列 $\mathbf{A}^-$ として，一般逆行列を定義する人もいるが，それらの条件は多くの応用分野で不必要である（Rao (1955 a, 1962 f)）．しかしながら，特殊の分野では，一般逆行列に適当な条件を付加することができ，そのような場合には，指定した性質をもつ一般逆行列としてそれを引用する．

（ii）<u>$\mathbf{A}^-$ は $\mathbf{A}$ の任意の一般逆行列で $\mathbf{A}^-\mathbf{A} = \mathbf{H}$ とする．このとき，</u>

（a）<u>$\mathbf{H}^2 = \mathbf{H}$，すなわち，$\mathbf{H}$ はベキ等である．</u>

（b）<u>$\mathbf{AH} = \mathbf{A}$ で $\operatorname{rank}\mathbf{A} = \operatorname{rank}\mathbf{H} = \operatorname{trace}\mathbf{H}$.</u>

（c）<u>$\mathbf{AX} = \mathbf{0}$ の一般解は，$(\mathbf{H} - \mathbf{I})\mathbf{Z}$ で，$\mathbf{Z}$ は任意である．</u>

（d）<u>解をもつ方程式 $\mathbf{AX} = \mathbf{Y}$ の一般解は，$\mathbf{A}^-\mathbf{Y} + (\mathbf{H} - \mathbf{I})\mathbf{Z}$ と書け，$\mathbf{Z}$ は任意で</u>

ある．

（e）$\mathbf{Q}'\mathbf{H}=\mathbf{Q}'$ が成り立つときのみ，$\mathbf{Q}'\mathbf{X}$ は，$\mathbf{AX}=\mathbf{Y}$ を満足するすべての $\mathbf{X}$ にたいして，一意な値をとる．

$\mathbf{AA}^-\mathbf{A}=\mathbf{A}$ であるから，$\mathbf{A}^-\mathbf{AA}^-\mathbf{A}=\mathbf{A}^-\mathbf{A}$，すなわち $\mathbf{H}^2=\mathbf{H}$ となる．他の結果を証明するのは容易である．

（iii）$\mathbf{A}^-$ は存在し，かつ，$\operatorname{rank}\mathbf{A}^-\geqslant\operatorname{rank}\mathbf{A}$ である．

[(iii), 1b. 2] において，与えられた行列 $\mathbf{A}$ にたいし，正則な正方行列 $\mathbf{B}$ と $\mathbf{C}$ が存在し，$\mathbf{BAC}=\Delta$，つまり，$\mathbf{A}=\mathbf{B}^{-1}\Delta\mathbf{C}^{-1}$ と書けることを示した．ここで，Δ は対角（必ずしも正方ではない）行列である．Δ^- を，Δ' の非零要素をその逆数でおきかえて得られる行列とする．すると，$\Delta\Delta^-\Delta=\Delta$ であることは容易にわかる．$\mathbf{A}^-=\mathbf{C}\Delta^-\mathbf{B}$ を考える．

$$\mathbf{AA}^-\mathbf{A}=\mathbf{B}^{-1}\Delta\mathbf{C}^{-1}\mathbf{C}\Delta^-\mathbf{B}\mathbf{B}^{-1}\Delta\mathbf{C}^{-1}=\mathbf{B}^{-1}\Delta\mathbf{C}^{-1}=\mathbf{A}$$

が証明できるから，$\mathbf{A}^-$ は一般逆行列である．明らかに，$\operatorname{rank}\mathbf{A}^-\geqslant\operatorname{rank}\mathbf{A}$ である．

実際，こうしてつくった $\mathbf{A}^-$ は，付加条件 $\mathbf{A}^-\mathbf{AA}^-=\mathbf{A}^-$ をみたし，よって，$(\mathbf{A}^-)^-=\mathbf{A}$ で，$\operatorname{rank}\mathbf{A}^-=\operatorname{rank}\mathbf{A}$ となる．しかし，(v) に示すように，すべての一般逆行列がこの付加条件をみたす必要はない．

（iv）$\mathbf{A}^-$ は一般逆行列である $\Leftrightarrow$ $\mathbf{A}^-\mathbf{A}$ がベキ等で $R(\mathbf{A}^-\mathbf{A})=R(\mathbf{A})$，あるいは，$\mathbf{AA}^-$ がベキ等で $R(\mathbf{AA}^-)=R(\mathbf{A})$ である．

この結果を証明するのも容易である．命題 (i) と (iv) は，一般逆行列の別の定義を与えている．（行列 $\mathbf{B}$ は，$\mathbf{B}^2=\mathbf{B}$ のとき，ベキ等といわれる．1b. 7 を見よ．）

（v）すべての一般逆行列の集合　$\mathbf{AXA}=\mathbf{A}$ の一般解（すなわち，$\mathbf{A}$ の一般逆行列）は，3つの異なる形であらわされる．

（a）$\mathbf{X}=\mathbf{A}^-+\mathbf{U}-\mathbf{A}^-\mathbf{AUAA}^-$　(1b. 5. 2)

（b）$\mathbf{X}=\mathbf{A}^-+\mathbf{V}(\mathbf{I}-\mathbf{AA}^-)+(\mathbf{I}-\mathbf{A}^-\mathbf{A})\mathbf{W}$　(1b. 5. 3)

（c）$\mathbf{X}=\mathbf{A}^-+\mathbf{V}(\mathbf{I}-\mathbf{P}_{\mathbf{A}})+(\mathbf{I}-\mathbf{P}_{\mathbf{A}^*})\mathbf{U}$　(1b. 5. 4)

ここで，$\mathbf{A}^-$ はある特定の一般逆行列，$\mathbf{P}_{\mathbf{A}}$ と $\mathbf{P}_{\mathbf{A}^*}$ はそれぞれ $\mathscr{M}(\mathbf{A})$ と $\mathscr{M}(\mathbf{A}^*)$ の上への射影子，$\mathbf{U},\mathbf{V},\mathbf{W}$ は任意の行列である．

(1b. 5. 2) の $\mathbf{X}$ は，明らかに，方程式 $\mathbf{AXA}=\mathbf{A}$ を満足する．さらに，任意の解 $\mathbf{X}$ は，$\mathbf{U}=\mathbf{X}-\mathbf{A}^-$ にとると，(1b. 5. 2) の形に書くことができる．同様に，(1b. 5. 3) の $\mathbf{X}$ も方程式 $\mathbf{AXA}=\mathbf{A}$ をみたす．さらに，任意の解 $\mathbf{X}$ は，$\mathbf{V}=\mathbf{X}-\mathbf{A}^-$，$\mathbf{W}=\mathbf{XAA}^-$ に選ぶと，(1b. 5. 3) の形に書ける．(1b. 5. 4) の結果も，同じ方法で証明できる．

定義　条件 $\mathbf{AGA}=\mathbf{A}$ と $\mathbf{GAG}=\mathbf{G}$ をみたす $\mathbf{A}$ の一般逆行列 $\mathbf{G}$ は，$\mathbf{A}$ の反射型 (reflexive) 一般逆行列と呼ばれ，$\mathbf{A}_r^-$ であらわす．この条件は，$\mathbf{AGA}=\mathbf{A}$ で $R(\mathbf{G})=R(\mathbf{A})$ と書いても同じである．

(iii) で与えた一般逆行列は，実は，反射型一般逆行列である．しかし，(1b. 5. 2) から (1b. 5. 4) までは，すべての一般逆行列が反射型である必要はないことを示す．

（vi）次の諸結果は，どのような一般逆行列にたいしても成立する．

（a）$\mathbf{A}(\mathbf{A}^*\mathbf{A})^-\mathbf{A}^*\mathbf{A}=\mathbf{A}$，かつ $\mathbf{A}^*\mathbf{A}(\mathbf{A}^*\mathbf{A})^-\mathbf{A}^*=\mathbf{A}^*$　(1b. 5. 5)

（b）$\mathbf{BA}^-\mathbf{A}=\mathbf{B}$ であるための必要十分条件は，$\mathscr{M}(\mathbf{B}')\subset\mathscr{M}(\mathbf{A}')$ である．

（c）$\mathbf{B}$ と $\mathbf{C}$ を非零行列とする．このとき，$\mathbf{BA}^-\mathbf{C}$ が $\mathbf{A}^-$ をどのように選んでも不変であるための必要十分条件は，$\mathscr{M}(\mathbf{B}')\subset\mathscr{M}(\mathbf{A}')$，$\mathscr{M}(\mathbf{C})\subset\mathscr{M}(\mathbf{A})$ である．

結果の (a) は $\mathbf{A}(\mathbf{A}^*\mathbf{A})^-\mathbf{A}^*\mathbf{A}-\mathbf{A}$ に，その共役転置行列を〔前から〕かけ，その積が零になることを示せば証明できる．(b), (c) の結果も同様に証明できる．

（vii）$\mathbf{A}(\mathbf{A}^*\mathbf{A})^-\mathbf{A}^*$ はエルミート行列で，一般逆行列 $(\mathbf{A}^*\mathbf{A})^-$ をどのように選んでも不変である．

差 $\mathbf{D}=\mathbf{A}(\mathbf{A}^*\mathbf{A})^-\mathbf{A}^*-\mathbf{A}[(\mathbf{A}^*\mathbf{A})^-]^*\mathbf{A}^*$ を考え，(vi) の (a) を用いて $\mathbf{DD}^*=\mathbf{0}$ となることを示せば，エルミート性を証明できる．不変性は $\mathscr{M}(\mathbf{A}^*)\subset\mathscr{M}(\mathbf{A}^*\mathbf{A})$ から出る．

（viii）ムーア＝ペンローズ逆行列　Moore (1935) と Penrose (1955) は，普通の一般逆行列 $\mathbf{A}^-$ と区別して，次の性質をみたす一般逆行列 $\mathbf{A}^+$ を定義した．

（a）$\mathbf{AA}^+\mathbf{A}=\mathbf{A}$，　（b）$\mathbf{A}^+\mathbf{AA}^+=\mathbf{A}^+$，

（c）$(\mathbf{AA}^+)^*=\mathbf{AA}^+$，　（d）$(\mathbf{A}^+\mathbf{A})^*=\mathbf{A}^+\mathbf{A}$．　(1b. 5. 6)

このような逆行列は存在し，一意的である．

1c. 3 の (1c. 3. 8) に示したように，$\mathbf{B}$ と $\mathbf{C}$ を直交行列，Δ を対角（必ずしも正方ではない）行列とすると，$\mathbf{BAC}=\Delta$ となるような $\mathbf{A}$ の対角化が存在する．すると，$\mathbf{A}^+=\mathbf{C}\Delta^-\mathbf{B}$ は条件 (a) から (d) までをみたし，かつ，(a) から (d) までをみたす他の行列を $\mathbf{D}$ とすると，$\mathbf{D}=\mathbf{A}^+$ となることを証明するのは容易である．

一般逆行列に関するこれ以上の結果は 1c. 5 節で述べる．

一般逆行列の若干の計算　$\mathbf{A}$ を階数 r の $m\times n$ 行列とする．このとき，必要なら，$\mathbf{A}$ のいくつかの行と列を除いて得られる，次数 r の正則な部分行列 $\mathbf{B}$ が存在する．列と行を適当に交換して，$\mathbf{A}$ を分割行列

$$\mathbf{A}=\begin{pmatrix}\mathbf{B} & \mathbf{C}\\ \mathbf{D} & \mathbf{E}\end{pmatrix}$$

の形に書くことができる．

$$\mathbf{A}^-=\begin{pmatrix}\mathbf{B}^{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}\qquad(1b.\ 5.\ 7)$$

を考える．$\mathbf{AA}^-\mathbf{A}=\mathbf{A}$ であり，したがって，(1b. 5. 7) で定義した $\mathbf{A}^-$ は一般逆行列で，$\operatorname{rank}\mathbf{A}^-=\operatorname{rank}\mathbf{A}$ を満足し，かつ，付加条件 $\mathbf{A}^-\mathbf{AA}^-=\mathbf{A}^-$ も成立することを，（$\mathbf{E}=$

$\mathbf{DB}^{-1}\mathbf{C}$ に注目すれば）証明することは容易である．

$\mathbf{BK}=\mathbf{C}$，$\mathbf{DK}=\mathbf{E}$ をみたすような $\mathbf{K}$ をとって，

$$\mathbf{A}^{-}=\begin{pmatrix}\mathbf{B}^{-1} & -\mathbf{K}\\ \mathbf{0} & \mathbf{I}\end{pmatrix} \tag{1b.5.8}$$

を考える．このとき $\operatorname{rank}\mathbf{A}^{-}=\min(m,n)$ となり，$\mathbf{A}^{-}$ は階数が最大の一般逆行列となる．

一般逆行列を計算する他の方法については，1b.8 の終わりの問題5を見よ．

1b.6　ベクトル空間，基底などの行列表現

これまで，行列の階数を，行列の列または行によって生成されるベクトル空間の次元として定義した以外には，ベクトル空間の概念を行列の研究に十分には利用してこなかった．ここで，ベクトル空間の論法を用いて，行列に関するいくつかの結果を証明しよう．それらはあとの章でも繰り返し用いられる．$\mathbf{X}$ を $n\times m$ 行列とする．$\mathbf{X}$ の列によって生成されるベクトル空間を $\mathscr{M}(\mathbf{X})$ であらわし，$\mathbf{X}$ の列空間とも呼ぶ．E_m で $\mathbf{XZ}=\mathbf{0}$ となるようなすべてのベクトル $\mathbf{Z}$ の張る部分空間を，$\mathbf{X}$ の零空間と呼び，$\mathscr{N}(\mathbf{X})$ であらわす．その次元，$d[\mathscr{N}(\mathbf{X})]$ を $\mathbf{X}$ の零化数とよぶ．

（i）$\mathscr{M}(\mathbf{X})=\mathscr{M}(\mathbf{XX}')$．すなわち，$\mathbf{X}$ の列と $\mathbf{XX}'$ の列によって生成されるベクトル空間は同じである．したがって，$d[\mathscr{M}(\mathbf{X})]=d[\mathscr{M}(\mathbf{XX}')]=R(\mathbf{X})=R(\mathbf{XX}')$．

α を $\alpha'\mathbf{X}=\mathbf{0}$ であるような列ベクトルとすると，$\alpha'\mathbf{XX}'=\mathbf{0}$ となることを確かめよ．反対に，$\alpha'\mathbf{XX}'=\mathbf{0}\Rightarrow\alpha'\mathbf{XX}'\alpha=0\Rightarrow\alpha'\mathbf{X}=\mathbf{0}$，したがって，$\mathbf{X}$ に直交するベクトルはいずれも $\mathbf{XX}'$ にも直交する．それゆえ，$\mathscr{M}(\mathbf{X})=\mathscr{M}(\mathbf{XX}')$ となる．さらに，任意の行列 $\mathbf{B}$ にたいして，$\mathscr{M}(\mathbf{B})\supset\mathscr{M}(\mathbf{X}')$ のときのみ，$\mathscr{M}(\mathbf{X})=\mathscr{M}(\mathbf{XB})=\mathscr{M}(\mathbf{XBB}')$ であることを証明するのはやさしい．

（ii）α と γ を $\alpha=\mathbf{X}\gamma$ であるようなベクトルとすると，$\alpha=\mathbf{XX}'\beta$ であるようなベクトル β が存在する．

この結果は，$\alpha\in\mathscr{M}(\mathbf{X})\Rightarrow\alpha\in\mathscr{M}(\mathbf{XX}')$ であるから，正しい．

（iii）$\mathbf{H}$ を所与の $k\times m$ 行列とするとき，ベクトル $\{\mathbf{X}\gamma$：ただし γ は $\mathbf{H}\gamma=\mathbf{0}$ という条件をみたす$\}$ の集合は，部分空間 $\mathscr{S}\subset\mathscr{M}(\mathbf{X})$ である．さらに，

$$d[\mathscr{S}]=\operatorname{rank}\begin{pmatrix}\mathbf{X}\\ \mathbf{H}\end{pmatrix}-\operatorname{rank}\mathbf{H} \tag{1b.6.1}$$

である．

γ が任意のとき，$\mathbf{X}\gamma$ というベクトル全体の集合は $\mathscr{M}(\mathbf{X})$ と同じであることにまず注目せよ．しかし，γ が指定のような制約を受けるときは，その部分空間だけが張られ，その次元は式 (1b.6.1) によって計算できる．

1b.7　ベキ等行列

定義　正方行列 $\mathbf{A}$ は，$\mathbf{A}^2=\mathbf{A}$ のとき，ベキ等と言う．

ベキ等行列に関する次の諸命題は重要である．

（i）ベキ等行列の階数は，そのトレース trace，つまり，その対角要素の和（対角和ともいう）に等しい．

$\mathbf{A}=\mathbf{B}_1\mathbf{C}_1$ という階数分解 (1b.2.3) を考える．このとき，$\mathbf{A}^2=\mathbf{A}\Rightarrow\mathbf{B}_1\mathbf{C}_1\mathbf{B}_1\mathbf{C}_1=\mathbf{B}_1\mathbf{C}_1$．ところで，$\mathbf{B}_1$ は左逆行列 $\mathbf{L}$ を，$\mathbf{C}_1$ は右逆行列 $\mathbf{R}$ をもつ．それゆえ，

$$\mathbf{LB}_1\mathbf{C}_1\mathbf{B}_1\mathbf{C}_1\mathbf{R}=\mathbf{LB}_1\mathbf{C}_1\mathbf{R}\Rightarrow\mathbf{C}_1\mathbf{B}_1=\mathbf{I}_r.$$

よって，$\operatorname{Trace}\mathbf{A}=\operatorname{Trace}\mathbf{B}_1\mathbf{C}_1=\operatorname{Trace}\mathbf{C}_1\mathbf{B}_1=\operatorname{Trace}\mathbf{I}_r=r=R(\mathbf{A})$．

（ii）$\mathbf{A}^2=\mathbf{A}\Leftrightarrow R(\mathbf{A})+R(\mathbf{I}-\mathbf{A})=m$，$m$ は $\mathbf{A}$ の次数．

証明は簡単である．

（iii）$\mathbf{A}_i$，$i=1,\cdots,k$ は m 次元の正方行列，$\mathbf{A}=\mathbf{A}_1+\cdots+\mathbf{A}_k$ とする．次の諸条件を考える．

（a）$\mathbf{A}_i^2=\mathbf{A}_i$，すべての i について．

（b）$\mathbf{A}_i\mathbf{A}_j=\mathbf{0}$，すべての $i\neq j$ について；$R(\mathbf{A}_i^2)=R(\mathbf{A}_i)$，すべての i について．

（c）$\mathbf{A}^2=\mathbf{A}$．

（d）$R(\mathbf{A})=R(\mathbf{A}_1)+\cdots+R(\mathbf{A}_k)$．

このとき，(a), (b), (c) のうちどれか2つがあれば，4つ全部が出てくる．さらに (c) と (d) $\Rightarrow$ (a) と (b)．

(a) と (b) $\Rightarrow$ (c) の証明は簡単．ベキ等行列の階数はそのトレースに等しいから，(a) と (c) $\Rightarrow$ (d)，したがって，(a) と (b) $\Rightarrow$ (c) と (d)．

(b) と (c) $\Rightarrow$ (a) を証明するには，(b) と (c) $\Rightarrow\mathbf{A}_i^2=\mathbf{AA}_i=\mathbf{A}^2\mathbf{A}_i=\mathbf{A}_i^3\Rightarrow\mathbf{A}_i^2(\mathbf{I}-\mathbf{A}_i)=\mathbf{0}$．また，ある $\mathbf{D}$ にたいして，$R(\mathbf{A}_i)=R(\mathbf{A}_i^2)\Rightarrow\mathbf{A}_i=\mathbf{DA}_i^2$．したがって，$\mathbf{A}_i^2(\mathbf{I}-\mathbf{A}_i)=\mathbf{0}\Rightarrow\mathbf{A}_i(\mathbf{I}-\mathbf{A}_i)=\mathbf{0}$．一方，(a) と (b) $\Rightarrow$ (c) と (d) はすでに示した．それゆえ，(b) と (c) $\Rightarrow$ (a) と (d)．

(c) と (d) $\Rightarrow$ (a) と (b) を証明するためには，$\mathbf{A}_0=\mathbf{I}-\mathbf{A}$ と定義する．(c) と (d) $\Rightarrow\mathbf{A}_0+\mathbf{A}_1+\cdots+\mathbf{A}_k=\mathbf{I}$，かつ，$R(\mathbf{A}_0)+R(\mathbf{A}_1)+\cdots+R(\mathbf{A}_k)=m$．さらに，$R(\mathbf{I}-\mathbf{A}_i)=R(\sum\mathbf{A}_i-\mathbf{A}_i)\leqslant\sum R(\mathbf{A}_i)-R(\mathbf{A}_i)=m-R(\mathbf{A}_i)$．一方，$R(\mathbf{A}_i)+R(\mathbf{I}-\mathbf{A}_i)\geqslant m$．したがって，$R(\mathbf{A}_i)+R(\mathbf{I}-\mathbf{A}_i)=m$ を得，(ii) の結果を用いると，すべての i について $\mathbf{A}_i^2=\mathbf{A}_i$ を得る．同様に，$(\mathbf{A}_i+\mathbf{A}_j)^2=\mathbf{A}_i+\mathbf{A}_j$．このとき，

$$\begin{aligned}(\mathbf{A}_i+\mathbf{A}_j)^2=\mathbf{A}_i+\mathbf{A}_j,\quad \mathbf{A}_i^2=\mathbf{A}_i,\quad \mathbf{A}_j^2=\mathbf{A}_j\\ \Rightarrow\mathbf{A}_i\mathbf{A}_j+\mathbf{A}_j\mathbf{A}_i=\mathbf{0}\\ \Rightarrow(\mathbf{A}_i\mathbf{A}_j+\mathbf{A}_j\mathbf{A}_i)\mathbf{A}_j=\mathbf{A}_i\mathbf{A}_j+\mathbf{A}_j\mathbf{A}_i\mathbf{A}_j=\mathbf{0}\end{aligned} \tag{1b.7.1}$$

$$\Rightarrow \mathbf{A}_j(\mathbf{A}_i\mathbf{A}_j + \mathbf{A}_j\mathbf{A}_i\mathbf{A}_j) = 2\mathbf{A}_j\mathbf{A}_i\mathbf{A}_j = \mathbf{0}. \qquad (1b.7.2)$$

(1b.7.1) に (1b.7.2) を代入すると，すべての $i \neq j$ について，$\mathbf{A}_i\mathbf{A}_j = \mathbf{0}$ を得る．この結果は Khatri (1968) による．

(iv) $\mathbf{G}_i$ を階数 r_i の $m \times p_i$ 行列，$i = 1, \cdots, k$ とする．$r_1 + \cdots + r_k = m$ ならば，次の条件は同等である：

(a) $\mathbf{G}_i^*\mathbf{G}_j = \mathbf{0}, \quad i \neq j$ (1b.7.3)

(b) $\mathbf{I} = \mathbf{G}_1(\mathbf{G}_1^*\mathbf{G}_1)^-\mathbf{G}_1^* + \cdots + \mathbf{G}_k(\mathbf{G}_k^*\mathbf{G}_k)^-\mathbf{G}_k^*$ (1b.7.4)

証明は，(iii) の (c) と (d) ⇒ (a) と (b) という結果を用いてできる．

1b.8　行列の特殊な積

(i) クロネッカー積　$\mathbf{A} = (a_{ij})$，$\mathbf{B} = (b_{ij})$ をそれぞれ $m \times n$，$p \times q$ の行列とする．このとき，クロネッカー積

$$\mathbf{A} \otimes \mathbf{B} = (a_{ij}\mathbf{B}) \qquad (1b.8.1)$$

は $mp \times nq$ 行列で，その (i, j) 番目の分割が $a_{ij}\mathbf{B}$ である分割行列としてあらわされる．ただし，$i = 1, \cdots, m$，$j = 1, \cdots, n$.

次の諸結果は，定義 (1b.8.1) から出る．

(a) $0 \otimes \mathbf{A} = \mathbf{A} \otimes 0 = 0$

(b) $(\mathbf{A}_1 + \mathbf{A}_2) \otimes \mathbf{B} = (\mathbf{A}_1 \otimes \mathbf{B}) + (\mathbf{A}_2 \otimes \mathbf{B})$

(c) $\mathbf{A} \otimes (\mathbf{B}_1 + \mathbf{B}_2) = (\mathbf{A} \otimes \mathbf{B}_1) + (\mathbf{A} \otimes \mathbf{B}_2)$

(d) $a\mathbf{A} \otimes b\mathbf{B} = ab\,\mathbf{A} \otimes \mathbf{B}$

(e) $\mathbf{A}_1\mathbf{A}_2 \otimes \mathbf{B}_1\mathbf{B}_2 = (\mathbf{A}_1 \otimes \mathbf{B}_1)(\mathbf{A}_2 \otimes \mathbf{B}_2)$

(f) $(\mathbf{A} \otimes \mathbf{B})^{-1} = \mathbf{A}^{-1} \otimes \mathbf{B}^{-1}$，これらの逆行列が存在するとき．

(g) 一般逆行列を用いて，$(\mathbf{A} \otimes \mathbf{B})^- = \mathbf{A}^- \otimes \mathbf{B}^-$

(h) $(\mathbf{A} \otimes \mathbf{B})' = \mathbf{A}' \otimes \mathbf{B}'$

(i) $(\mathbf{A} \otimes \mathbf{B})(\mathbf{A}^{-1} \otimes \mathbf{B}^{-1}) = \mathbf{I}$

(ii) アダマール積　$\mathbf{A} = (a_{ij})$ と $\mathbf{B} = (b_{ij})$ をともに $m \times n$ 行列とすると，そのアダマール積は，要素同士の積の $m \times n$ 行列である．

$$\mathbf{A} * \mathbf{B} = (a_{ij}b_{ij}). \qquad (1b.8.2)$$

アダマール積を含む次の結果が得られる．

(a) $\mathbf{A} * \mathbf{0} = \mathbf{0}$

(b) $\mathbf{A} * \mathbf{e}\mathbf{e}' = \mathbf{A} = \mathbf{e}\mathbf{e}' * \mathbf{A}$，ここに $\mathbf{e}' = (1, \cdots, 1)$.

(c) $\mathbf{A} * \mathbf{B} = \mathbf{B} * \mathbf{A}$

(d) $(\mathbf{A} + \mathbf{B}) * \mathbf{C} = \mathbf{A} * \mathbf{C} + \mathbf{B} * \mathbf{C}$

(e) $\mathrm{tr}\,\mathbf{AB} = \mathbf{e}'(\mathbf{A} * \mathbf{B}')\mathbf{e}$.

(iii) 新しい積　$\mathbf{A} = (\mathbf{A}_1 \vdots \cdots \vdots \mathbf{A}_k)$ と $\mathbf{B} = (\mathbf{B}_1 \vdots \cdots \vdots \mathbf{B}_k)$ は，同数の分割をもつ，2つの分割行列とする．Khatri and Rao (1968) は，積

$$\mathbf{A} \odot \mathbf{B} = (\mathbf{A}_1 \otimes \mathbf{B}_1 \vdots \cdots \vdots \mathbf{A}_k \otimes \mathbf{B}_k) \qquad (1b.8.3)$$

を定義した．ここで，⊗はクロネッカー積をあらわす．ここで，

$$(\mathbf{A} \odot \mathbf{B}) \odot \mathbf{C} = \mathbf{A} \odot (\mathbf{B} \odot \mathbf{C}), \qquad (\mathbf{T}_1 \otimes \mathbf{T}_2)(\mathbf{A} \odot \mathbf{B}) = \mathbf{T}_1\mathbf{A} \odot \mathbf{T}_2\mathbf{B}$$

などを証明するのは容易である．

補足と問題

1　行列の階数　$R(\mathbf{A})$ で $\mathbf{A}$ の階数をあらわす．

1.1　$R(\mathbf{AB}) \leqslant \min[R(\mathbf{A}), R(\mathbf{B})]$

[$\mathbf{AB}$ の行は $\mathbf{B}$ の行の1次結合であるから，$\mathbf{AB}$ の独立な行の数は，$\mathbf{B}$ の独立な行の数より少ないか等しい．したがって，$R(\mathbf{AB}) \leqslant R(\mathbf{B})$．同様に，もう1つの不等式も得られる．]

1.2　(a) $R(\mathbf{AA}') = R(\mathbf{A})$

(b) $\mathbf{B}$ が正則な正方行列であるなら，$R(\mathbf{AB}) = R(\mathbf{A})$.

[問題1.1によって，$R(\mathbf{AB}) \leqslant R(\mathbf{A})$．$\mathbf{A} = (\mathbf{AB})(\mathbf{B}^{-1})$．よって，$R(\mathbf{A}) = R[(\mathbf{AB})(\mathbf{B}^{-1})] \leqslant R(\mathbf{AB})$．したがって，$R(\mathbf{AB}) = R(\mathbf{A})$.]

(c) $\mathbf{A}$ を階数 m の $m \times n$ 行列，$\mathbf{S}$ を階数 r の $r \times m$ 行列とする．このとき $R(\mathbf{SA}) = r$.

1.3　$R(\mathbf{A}) = s$ なら s より大きい次数の小行列式はすべて零になり，反対に，s より大きい次数の小行列式はすべて零で，次数 s の小行列式のなかに少なくとも1つ零でないものがあれば，$R(\mathbf{A}) = s$ である．

1.4　シルベスターの法則　$\mathbf{A}$ と $\mathbf{B}$ が次数 n の正方行列で，$R(\mathbf{A}) = r$，$R(\mathbf{B}) = s$ であるなら，$R(\mathbf{AB}) \geqslant r + s - n$.

1.5　階数 r の行列はいずれも，各々階数1の r 個の行列の和として書ける．

1.6　フロベニウスの不等式　$\mathbf{A}, \mathbf{B}, \mathbf{C}$ を以下に考える積が定義できるような3つの行列とすると，

$$R(\mathbf{AB}) + R(\mathbf{BC}) \leqslant R(\mathbf{B}) + R(\mathbf{ABC}).$$

ここで，$\mathbf{B} = \mathbf{I}$ (次数 $n \times n$) とおくと，$\mathbf{A}$ は $m \times n$，$\mathbf{C}$ は $n \times p$ となり，

$$R(\mathbf{A}) + R(\mathbf{C}) - R(\mathbf{AC}) \leqslant n$$

を得る．また，$\mathbf{AC} = \mathbf{0}$ なら，$R(\mathbf{A}) + R(\mathbf{C}) \leqslant n$ となる．

1.7　$\mathbf{A} - \mathbf{I}$ と $\mathbf{B} - \mathbf{I}$ の階数を p, q とすると，$\mathbf{AB} - \mathbf{I}$ の階数は $p + q$ 以下となる．

1.8　$\mathbf{A}$ と $\mathbf{B}$ を $\mathbf{A} + \mathbf{B}$ が定義できるような2つの行列とすると，$R(\mathbf{A} + \mathbf{B}) \leqslant R(\mathbf{A}) + R(\mathbf{B})$ となる．$R(\mathbf{A} + \mathbf{B}) = R(\mathbf{A}) + R(\mathbf{B}) \Leftrightarrow \mathscr{M}(\mathbf{A}) \cap \mathscr{M}(\mathbf{B}) = \{\mathbf{0}\}$ で，かつ，$\mathscr{M}(\mathbf{A}') \cap \mathscr{M}(\mathbf{B}') = \{\mathbf{0}\}$.

1.9　次の諸条件は同等である．

(a) $R(\mathbf{A}) = R(\mathbf{A}^2)$.

（b）　$\mathscr{M}(\mathbf{A}) \cap \mathscr{N}(\mathbf{A}) = \{0\}$，ただし，$\mathscr{N}(\mathbf{A})$ は $\mathbf{A}$ の零空間.

（c）　$\mathbf{D}$ を正則とするとき，

$$\mathbf{A} = \mathbf{P}\begin{pmatrix} \mathbf{D} & 0 \\ 0 & 0 \end{pmatrix}\mathbf{P}^{-1}$$

とするような正則行列 $\mathbf{P}$ が存在する.

1.10　$R(\mathbf{A}^m) = R(\mathbf{A}^{m+1}) \Rightarrow R(\mathbf{A}^m) = R(\mathbf{A}^n)$，すべての $n \geqq m$ にたいして.

1.11　任意の正方行列 $\mathbf{A}$ にたいして，$R(\mathbf{A}^m) = R(\mathbf{A}^{m+1})$ となるような正の整数 m が存在する.（そのような最小の正の整数を $\mathbf{A}$ の指標とよぶ.）

2　行列式

2.1　次数 n の正方行列 $\mathbf{A} = (a_{ij})$ は，$a_{ii} = 0$ かつ $a_{ij} = -a_{ji}$ のとき，歪対称 skew symmetric であるという．次を証明せよ.

（a）　n が奇数なら $|\mathbf{A}| = 0$，n が偶数なら完全平方の形となる.

（b）　$R(\mathbf{A})$ はつねに偶数である.

（c）　どの行列も，対称行列と歪対称行列との和の形に書くことができる.

2.2　$\mathbf{A}$ を $m \times n$ とするとき，

（a）　$m \leqq n$ のとき $|\mathbf{AA}'| \geqq 0$，$m > n$ のとき $= 0$，$R(\mathbf{A}) = m$ なら > 0 である.

（b）　$m \leqq n$ のとき $|\mathbf{AA}'|$ は $\mathbf{A}$ の m 列でできるすべての行列式の2乗和に等しい.

2.3　次で定義されるファンデルモンド行列式は，$\prod_{i<j}(d_j - d_i)$ という値をもつ：

$$\begin{pmatrix} 1 & 1 & \cdots & 1 \\ d_1 & d_2 & \cdots & d_n \\ \vdots & \vdots & & \vdots \\ d_1^{n-1} & d_2^{n-1} & \cdots & d_n^{n-1} \end{pmatrix}$$

2.4　$\mathbf{A}, \mathbf{D}$ を正方行列（次数は異なってもよい），$\mathbf{A}$ を正則とするとき，

$$\begin{vmatrix} \mathbf{A} & \mathbf{C} \\ \mathbf{B} & \mathbf{D} \end{vmatrix} = |\mathbf{A}|\,|\mathbf{D} - \mathbf{BA}^{-1}\mathbf{C}|$$

と書けることを示せ．この結果は，値1をもつ次の行列式を前からかければ得られる.

$$\begin{vmatrix} \mathbf{I} & 0 \\ -\mathbf{BA}^{-1} & \mathbf{I} \end{vmatrix}\begin{vmatrix} \mathbf{A} & \mathbf{C} \\ \mathbf{B} & \mathbf{D} \end{vmatrix} = \begin{vmatrix} \mathbf{A} & \mathbf{C} \\ 0 & \mathbf{D} - \mathbf{BA}^{-1}\mathbf{C} \end{vmatrix} = |\mathbf{A}|\,|\mathbf{D} - \mathbf{BA}^{-1}\mathbf{C}|.$$

ここで，次を確かめよ.

$$\begin{vmatrix} \mathbf{E} & \mathbf{F} \\ 0 & \mathbf{G} \end{vmatrix} = |\mathbf{E}|\,|\mathbf{G}|.$$

2.5　縁つき行列式の展開　　$\mathbf{A}$ を $n \times n$，$\mathbf{X}$ を $n \times 1$ 行列とする．さらに，A_{ij} を a_{ij} の余因子として，$\mathrm{adj}\,\mathbf{A} = (A_{ij})'$ と書く．このとき，

（a）　$\begin{vmatrix} \mathbf{A} & \mathbf{X} \\ \mathbf{X}' & c \end{vmatrix} = c|\mathbf{A}| - \mathbf{X}'(\mathrm{adj}\,\mathbf{A})\mathbf{X}$

（b）　$\begin{vmatrix} \mathbf{A} & \mathbf{X} \\ \mathbf{X}' & -1 \end{vmatrix} = -|\mathbf{A} + \mathbf{XX}'| = -|\mathbf{A}| - \mathbf{X}'(\mathrm{adj}\,\mathbf{A})\mathbf{X}$

$$= -|\mathbf{A}|(1 + \mathbf{X}'\mathbf{A}^{-1}\mathbf{X}) \qquad (|\mathbf{A}| \neq 0 \text{ ならば})$$

（c）　$\mathbf{U}$ を要素がすべて1の行列，$\mathbf{V}$ を成分がすべて1の列ベクトルとすると，

$$|\mathbf{A} + c\mathbf{U}| = |\mathbf{A}| + c\mathbf{V}'(\mathrm{adj}\,\mathbf{A})\mathbf{V}$$

$$\mathbf{V}'\mathrm{adj}(\mathbf{A} + c\mathbf{U}) = \mathbf{V}'(\mathrm{adj}\,\mathbf{A})$$
$$[\mathrm{adj}(\mathbf{A} + c\mathbf{U})]\mathbf{V} = (\mathrm{adj}\,\mathbf{A})\mathbf{V}.$$

2.6　長方形行列の積の行列式　　$\mathbf{A}$ を $m \times n$，$\mathbf{B}$ を $n \times m$ とし，$m \leqq n$ と仮定する．このとき，

$$|\mathbf{AB}| = \sum |\mathbf{A}_m|\,|\mathbf{B}_m|.$$

ここで，和は，次数 m の部分行列の全部についてとる．この $\mathbf{B}_m$ は，$\mathbf{B}$ の行からつくられる次数 m の正方部分行列で，$\mathbf{A}_m$ を規定する $\mathbf{A}$ の列と同じ指標をもつものである.

2.7　$\mathbf{A}, \mathbf{D}$ を次の式に現われる逆行列が存在するような対称行列とするとき，

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}' & \mathbf{D} \end{pmatrix}^{-1} = \begin{pmatrix} \mathbf{A}^{-1} + \mathbf{FE}^{-1}\mathbf{F}' & -\mathbf{FE}^{-1} \\ -\mathbf{E}^{-1}\mathbf{F}' & \mathbf{E}^{-1} \end{pmatrix}$$

となることを証明せよ．ここで，$\mathbf{E} = \mathbf{D} - \mathbf{B}'\mathbf{A}^{-1}\mathbf{B}$，$\mathbf{F} = \mathbf{A}^{-1}\mathbf{B}$ である．この問題によって，次数の高い行列の逆行列が，次数のより低い行列の逆行列だけを求めることによって計算する方法が得られる．1つの興味あるケースは，$\mathbf{B}$ が列ベクトルのときで，次数 $n+1$ の行列の逆が次数 n の主小行列の逆が与えられると計算できるという方法となる．この問題で導いた式を用いて，次の行列の逆を求めよ.

$$\begin{pmatrix} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & 4 \\ 0 & 0 & 1 & 2 \\ 3 & 4 & 2 & 5 \end{pmatrix}$$

［ヒント：　$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \mathbf{A}^{-1}$

$\mathbf{B}' = (3 \quad 4 \quad 2)$，$\mathbf{D} = (5)$ にとれ.］

2.8　$\mathbf{A}$ を正則行列，$\mathbf{U}$ と $\mathbf{V}$ を2つの列ベクトルとする．このとき，

$$(\mathbf{A} + \mathbf{UV}')^{-1} = \mathbf{A}^{-1} - \frac{(\mathbf{A}^{-1}\mathbf{U})(\mathbf{V}'\mathbf{A}^{-1})}{1 + \mathbf{V}'\mathbf{A}^{-1}\mathbf{U}}$$

となる．これは，$\mathbf{A}$ の逆を知って，$\mathbf{A} + \mathbf{UV}'$ の逆を計算する方法を与える.

2.9　$\mathbf{A}$ と $\mathbf{D}$ を次数 m, n の正則行列，$\mathbf{B}$ を $m \times n$ 行列とする．このとき，

$$(\mathbf{A} + \mathbf{BDB}')^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}(\mathbf{B}'\mathbf{A}^{-1}\mathbf{B} + \mathbf{D}^{-1})^{-1}\mathbf{B}'\mathbf{A}^{-1}$$
$$= \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{BEB}'\mathbf{A}^{-1} + \mathbf{A}^{-1}\mathbf{BE}(\mathbf{E} + \mathbf{D})^{-1}\mathbf{EB}'\mathbf{A}^{-1}$$

となる．ただし，$\mathbf{E} = (\mathbf{B}'\mathbf{A}^{-1}\mathbf{B})^{-1}$ である.

3　行列の対角和

　　正方行列 $\mathbf{A}$ のトレース（対角和）とは，その対角要素の和であって，$\mathrm{Tr}\,\mathbf{A}$ と書かれる．次を証明せよ，

（a）　$\mathrm{Tr}(\mathbf{A} + \mathbf{B}) = \mathrm{Tr}\,\mathbf{A} + \mathrm{Tr}\,\mathbf{B}$

（b）　$\mathrm{Tr}\,\mathbf{AB} = \mathrm{Tr}\,\mathbf{BA}$

（c）　$\mathrm{Tr}\,\mathbf{S}^{-1}\mathbf{AS} = \mathrm{Tr}\,\mathbf{A}$

（d）　$\mathbf{A}$ がベキ等なら，$\mathrm{Tr}\,\mathbf{A} = R(\mathbf{A})$

（e）　$\mathbf{A}^-$ を $\mathbf{A}$ の一般逆行列とすると，$\mathrm{Tr}(\mathbf{AA}^-) = R(\mathbf{A})$.

（f）　$\mathbf{X}$ を列ベクトルとすると，$\mathbf{X}'\mathbf{AX} = \mathrm{Tr}\,\mathbf{AXX}'$.

4　数理統計学で有用な変換行列の構成

4.1　$\mathbf{A}$ を階数 r の $n \times m$ 行列とする．分割行列 $\mathbf{B}_1$ は $n \times r$ で $\mathscr{M}(\mathbf{B}_1) = \mathscr{M}(\mathbf{A})$ であるような，次数 $n \times n$ の直交行列 $(\mathbf{B}_1 \vdots \mathbf{B}_2)$ が存在することを示せ.

4.2　$\mathbf{A}_1$ を階数 r の $n \times m$ 行列，$\mathbf{A}_2$ を階数 s の $n \times k$ 行列とし，$\mathscr{M}(\mathbf{A}_1) \cap \mathscr{M}(\mathbf{A}_2)$ は t 次元であるとする．このとき，直交行列 $(\mathbf{B}_1 \vdots \mathbf{B}_2 \vdots \mathbf{B}_3)$ が存在して，$\mathbf{B}_1$ は $n \times r$ で，$\mathscr{M}(\mathbf{B}_1) = \mathscr{M}(\mathbf{A}_1)$，$\mathbf{B}_2$ は $n \times (s-t)$ で $\mathscr{M}(\mathbf{B}_2) \subset \mathscr{M}(\mathbf{A}_2)$ とできる.

5　一般逆行列の計算　　$\mathbf{A}$ を階数 r の $n\times m$ 行列とする．このとき，次数 $s\times m$ で階数 $m-r$ の行列 $\mathbf{B}$ が存在して，$\mathscr{M}(\mathbf{A}')\cap\mathscr{M}(\mathbf{B}')$ が零ベクトルであるようにできる．次を示せ．

（a）$\mathbf{A}'\mathbf{A}+\mathbf{B}'\mathbf{B}$ の階数は m で，かつ，

（b）$\mathbf{A}'\mathbf{A}(\mathbf{A}'\mathbf{A}+\mathbf{B}'\mathbf{B})^{-1}\mathbf{A}'\mathbf{A}=\mathbf{A}'\mathbf{A}$

となり，したがって，$(\mathbf{A}'\mathbf{A}+\mathbf{B}'\mathbf{B})^{-1}$ は事実上，$\mathbf{A}'\mathbf{A}$ の一般逆行列となる．

（c）$s=m-r$ なら，

$$\begin{pmatrix}\mathbf{A}\mathbf{A}' & \mathbf{B}'\\ \mathbf{B} & 0\end{pmatrix}\quad \text{は，最大階数をもち，かつ}$$

（d）$\mathbf{C}_1$ を次のように定義すると，$\mathbf{C}_1$ は $\mathbf{A}'\mathbf{A}$ の一般逆行列となる．

$$\begin{pmatrix}\mathbf{A}'\mathbf{A} & \mathbf{B}'\\ \mathbf{B} & 0\end{pmatrix}^{-1}=\begin{pmatrix}\mathbf{C}_1 & \mathbf{C}_2'\\ \mathbf{C}_2 & \mathbf{C}_3\end{pmatrix}.$$

6　$\mathbf{A}$ と $\mathbf{B}$ を同数の列をもつ行列とし，$\mathscr{M}(\mathbf{A}')\cap\mathscr{M}(\mathbf{B}')=\{0\}$ とする．このとき，

（a）$R(\mathbf{A}'\mathbf{A}+\mathbf{B}'\mathbf{B})=R(\mathbf{A})+R(\mathbf{B})$，かつ

（b）一般逆行列をどのように選んでも，$\mathbf{A}'\mathbf{A}(\mathbf{A}'\mathbf{A}+\mathbf{B}'\mathbf{B})^{-}\mathbf{A}'\mathbf{A}=\mathbf{A}'\mathbf{A}$ となる．

7　$\mathbf{A}\mathbf{B}'=\mathbf{C}\mathbf{B}'\Leftrightarrow\mathbf{A}\mathbf{B}=\mathbf{C}\mathbf{B}$.

1c　固有値と行列の標準化

1c.1　2次形式の分類と変換

n 変数 $x_1,\cdots,x_n$ の2次形式とは，これらの変数の同次2次関数，

$$Q=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j=\mathbf{X}'\mathbf{A}\mathbf{X}$$

のことをいう．ここで，$\mathbf{X}$ は変数の列ベクトルであり，$\mathbf{A}=[(a_{ij}+a_{ji})/2]$ は対称で，2次形式の行列とよばれる．

2次形式の分類　　実2次形式 $\mathbf{X}'\mathbf{A}\mathbf{X}$ は，（a）すべての非零の $\mathbf{X}$ にたいして，$\mathbf{X}'\mathbf{A}\mathbf{X}>0$ なら，正定符号 p. d. (positive definite)，$-\mathbf{X}'\mathbf{A}\mathbf{X}$ が p. d. なら負定符号 n. d. (negative definite) とよばれる．また，（b）$\mathbf{X}'\mathbf{A}\mathbf{X}\geqq 0$ で，少なくとも1つの非零の $\mathbf{X}$ にたいして $\mathbf{X}'\mathbf{A}\mathbf{X}=0$ なら，半正定符号 p. s. d. (positive semidefinite)，$-\mathbf{X}'\mathbf{A}\mathbf{X}$ が p. s. d なら，半負定符号 n. s. d. (negative semidefinite) とよばれる．

p. d. あるいは p. s. d. の2次形式を非負定符号 n. n. d. (non-negative definite) と呼んだり，n. d. または n. s. d. の2次形式を非正定符号 n. p. d. と呼んだりするのも有用である．行列 $\mathbf{A}$ をもつ2次形式 $\mathbf{X}'\mathbf{A}\mathbf{X}$ が p. d.，n. d. などのとき，$\mathbf{A}$ 自身を p. d.，n. d. などと呼んで引用するのも習慣的である．

2次形式の変換　　正則な1次変換 $\mathbf{Y}=\mathbf{B}^{-1}\mathbf{X}$，つまり $\mathbf{X}=\mathbf{B}\mathbf{Y}$ を考える．このとき，

$$Q=\mathbf{X}'\mathbf{A}\mathbf{X}\to\mathbf{Y}'\mathbf{B}'\mathbf{A}\mathbf{B}\mathbf{Y}=\mathbf{Y}'\mathbf{C}\mathbf{Y}$$

となり，これは，行列 $\mathbf{C}=\mathbf{B}'\mathbf{A}\mathbf{B}$ をもつ新しい変数の2次形式となる．1次変換の下では，次の諸結果が成り立つ．

（i）正則な1次変換の下では，Q の符号は不変である．

ある $\mathbf{X}$ にたいして $\mathbf{X}'\mathbf{A}\mathbf{X}=g$（所与の実数）であるとき，$\mathbf{X}$ が非零なら $\mathbf{Y}$ も非零になるような変換 $\mathbf{Y}=\mathbf{B}^{-1}\mathbf{X}$ によって，

$$\mathbf{Y}'\mathbf{C}\mathbf{Y}=\mathbf{X}'\mathbf{B}^{-1\prime}\mathbf{B}'\mathbf{A}\mathbf{B}\mathbf{B}^{-1}\mathbf{X}=\mathbf{X}'\mathbf{A}\mathbf{X}=g$$

を得，反対に，ある $\mathbf{Y}$ にたいして $\mathbf{Y}'\mathbf{C}\mathbf{Y}=g$ のとき，$\mathbf{X}'\mathbf{A}\mathbf{X}=g$ を満足する $\mathbf{X}$ が存在する．それゆえ，非零の $\mathbf{X}$ にたいして $\mathbf{X}'\mathbf{A}\mathbf{X}$ のとる値は，非零の $\mathbf{Y}$ にたいして $\mathbf{Y}'\mathbf{C}\mathbf{Y}$ のとる値と同じである．これによって，符号の不変性を証明するのに必要なことより以上のことを証明した．

（ii）正則な1次変換によって，どの Q も2乗項だけを含む2次形式に標準化できる．

[(vi), 1b.2] において，$|\mathbf{B}|\neq 0$ で $\mathbf{B}'\mathbf{A}\mathbf{B}=\Delta$（対角行列）とできる行列 $\mathbf{B}$ が存在することを示した．そのような $\mathbf{B}$ を選んで，$\mathbf{Y}=\mathbf{B}^{-1}\mathbf{X}$ すなわち $\mathbf{X}=\mathbf{B}\mathbf{Y}$ という変換をする．すると，

$$\mathbf{X}'\mathbf{A}\mathbf{X}\to\mathbf{Y}'\mathbf{B}'\mathbf{A}\mathbf{B}\mathbf{Y}=\mathbf{Y}'\Delta\mathbf{Y}=d_1y_1^2+\cdots+d_ny_n^2\qquad(1c.1.1)$$

を得る．ここで，$d_1,\cdots,d_n$ は Δ の対角要素である．

$R(\mathbf{A})=r$ とする．$R(\mathbf{A})=R(\Delta)$ であるから，(1c.1.1) で零でない d_i の数は r に等しい．p と q を正の d_i と負の d_i の数とする．Q を2乗項だけを含む形式に標準化する任意の正則変換にたいして，p,q の値，したがって，2次形式の符号指数と呼ばれる $p-q$ の値は不変であることは容易に証明できる．明らかに，$p+q=r$ で，この値もすべての標準化について同じである．

一般性を失うことなく，$d_1,\cdots,d_p$ を正とできる．それに続く変換，$u_i=\sqrt{d_i}\,y_i$，$i=1,\cdots,p$；$u_i=\sqrt{-d_i}\,y_i$，$i=p+1,\cdots,r$；$u_i=y_i$，$i=r+1,\cdots,n$ を施すことによって，(1c.1.1) あるいは，$\mathbf{X}$ から $\mathbf{U}$ への直接の変換によって，もとの形式 $\mathbf{X}'\mathbf{A}\mathbf{X}$ は，

$$u_1^2+\cdots+u_p^2-u_{p+1}^2-\cdots-u_r^2\qquad(1c.1.2)$$

という形式に標準化される．ここで，p と r は，2次形式 $\mathbf{X}'\mathbf{A}\mathbf{X}$ に付与される不変な数である．

（iii）$\mathbf{X}'\mathbf{A}\mathbf{X}$ が p. d. なら，正則変換 $\mathbf{X}=\mathbf{B}\mathbf{Y}$ が存在して，

$$\mathbf{X}'\mathbf{A}\mathbf{X}\to\mathbf{Y}'\mathbf{Y}=y_1^2+\cdots+y_n^2\qquad(1c.1.3)$$

とできる．

符号は保持されるから，標準化した形式 (1c.1.1) のすべての項は正の係数 d_i をもち，したがって，それに続く変換 $\sqrt{d_i}\,y_i=u_i$ によって，形式 (1c.1.3) が得られる．$\mathbf{X}=$

$\mathbf{BY}$ をそのような変換とすると，

$$\mathbf{X'AX} \to \mathbf{Y'B'ABY} = \mathbf{Y'Y} \Rightarrow \mathbf{B'AB} = \mathbf{I},\ \text{すなわち}\ \mathbf{A} = \mathbf{CC'} \qquad (1c.1.4)$$

となり，$\mathbf{C'} = \mathbf{B}^{-1}$ である．これは，p. d. 行列 $\mathbf{A}$ が1つの行列とその転置行列との積としてあらわされることと同等である．もし $\mathbf{A}$ が n. n. d. でその階数 $r < n$ ならば，$\mathbf{X'AX} \to y_1^2 + \cdots + y_r^2$ とするような正則変換 $\mathbf{X} = \mathbf{BY}$ が存在することを，読者は証明せよ．(1c.1.4) の場合と同様に，階数 r の n. n. d. 行列 $\mathbf{A}$ は，階数 r の $n \times r$ 行列 $\mathbf{C}$ を用いて，$\mathbf{CC'}$ と書ける．

（iv）実2次形式 $\mathbf{X'AX}$ が p. d. であるための必要十分条件は，

$$g_i = \begin{vmatrix} a_{11} & \cdots & a_{1i} \\ & \cdots\cdots\cdots & \\ a_{i1} & \cdots & a_{ii} \end{vmatrix}$$

とおくと，$i = 1, \cdots, n$ にたいして $g_i > 0$ であることである．

（iii）によれば，

$$\mathbf{X'AX} \Leftrightarrow \mathbf{Y'B'ABY} = \mathbf{Y'Y}$$

とするような正則変換が存在し，$\mathbf{B'AB} = \mathbf{I}$，$|\mathbf{B'AB}| = |\mathbf{B}|^2|\mathbf{A}| = |\mathbf{I}| = 1$ となる．したがって，$|\mathbf{B}|^2 g_n > 0$ となり，よって，$g_n > 0$ を得る．次に，最後の変数 x_n が零であるような2次形式を考えると，これは，$n-1$ 変数について p. d. となる．したがって，これに上の議論を適用すると，その行列の行列式 $g_{n-1} > 0$ となる．以下同様にして，上の必要性は証明される．

$a_{11} > 0$ であるから，基本行列を前と後からかけることによって，$\mathbf{A}$ の第1列と第1行を，対称に掃き出すことができる．得られた行列は

$$\begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & b_{22} & \cdots & b_{2n} \\ & \cdots\cdots\cdots & & \\ 0 & b_{n2} & \cdots & b_{nn} \end{pmatrix}$$

となり，第1行第1列を含む小行列式はいずれもその値が変わらないから，$a_{11}b_{22} = g_2$ を得る．$g_2 > 0$ であるから，$b_{22} = g_2/g_1 > 0$ となる．b_{22} を枢軸として第2行第2列を掃き出すと，

$$\begin{pmatrix} a_{11} & 0 & 0 & \cdots & 0 \\ 0 & b_{22} & 0 & \cdots & 0 \\ \vdots & 0 & c_{33} & \cdots & c_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & c_{n3} & \cdots & c_{nn} \end{pmatrix}$$

を得，$a_{11}b_{22}c_{33} = g_3 > 0$ を得る．したがって，$c_{33} = g_3/g_2 > 0$ となる．この手順をつづけて行くと，最後には，

$$\begin{pmatrix} g_1 & 0 & \cdots & 0 \\ 0 & g_2/g_1 & \cdots & 0 \\ & \cdots\cdots\cdots & & \\ 0 & 0 & \cdots & g_n/g_{n-1} \end{pmatrix} \qquad (1c.1.5)$$

という形式になり，各対角要素は正となる．こうして，列の掃き出しに用いた基本行列の積である行列 $\mathbf{B}$ の存在と，それが $\mathbf{B'AB} = \Delta$ [(1c.1.5) という特殊の形をもつ] を満足することが証明された．したがって，$\mathbf{X} = \mathbf{BY}$ によって，$\mathbf{X'AX}$ は，

$$g_1 y_1^2 + \frac{g_2}{g_1} y_2^2 + \cdots + \frac{g_n}{g_{n-1}} y_n^2 \qquad (1c.1.6)$$

に変換され，正定符号である．この式は，十分性を証明している．また，$\mathbf{X'AX}$ が負定符号であるための必要十分条件は $g_1 < 0$，$g_2 > 0$，$g_3 < 0, \cdots$ であることも，2次形式 $-\mathbf{X'AX}$ を考えることからすぐわかる．

（iv）では，p. d. 2次形式 Q が，正則変換によって平方和に標準化できることを見た．次には，p. d. Q を新しい変数 $y_1, \cdots, y_n$ の互いに排反する部分集合の p. d. 2次形式の和に変換する変換 $\mathbf{Y} = \mathbf{CX}$ の条件を調べよう．このために $\mathbf{Y}$ を分割形式 $\mathbf{Y'} = (\mathbf{Y}_1' \vdots \cdots \vdots \mathbf{Y}_k')$ に書き，$\mathbf{Y}_i$ は，変数 $y_1, \cdots, y_n$ の i 番目の排反部分集合の列ベクトルとする．このとき，変換 $\mathbf{Y} = \mathbf{CX}$ は，$\mathbf{C}_1, \cdots, \mathbf{C}_k$ を $\mathbf{C}$ の分割，すなわち $\mathbf{C'} = (\mathbf{C}_1' \vdots \cdots \vdots \mathbf{C}_k')$ とすると，$\mathbf{Y}_1 = \mathbf{C}_1\mathbf{X}, \cdots, \mathbf{Y}_k = \mathbf{C}_k\mathbf{X}$ と書ける．すると，次の命題が成立する．

（v）$\mathbf{A}$ が p. d. のとき，正則変換 $\mathbf{Y}_i = \mathbf{C}_i\mathbf{X}$，$i = 1, \cdots, k$ の下で，

$$\mathbf{X'AX} \to \mathbf{Y}_1'\mathbf{B}_1\mathbf{Y}_1 + \cdots + \mathbf{Y}_k'\mathbf{B}_k\mathbf{Y}_k \qquad (1c.1.7)$$

となるための必要十分条件は $\mathbf{C}_i\mathbf{A}^{-1}\mathbf{C}_j' = 0$，$i \neq j$ で，この場合 $\mathbf{B}_i = (\mathbf{C}_i\mathbf{A}^{-1}\mathbf{C}_i')^{-1}$ となる．

1b.7 の (1b.7.3) と (1b.7.4) で，$\mathbf{A}^{-1/2}\mathbf{C}_i' = \mathbf{G}_i$ と書くと，次の条件

（a）$\mathbf{C}_i\mathbf{A}^{-1}\mathbf{C}_j' = 0$，すべての $i \neq j$ にたいして　(1c.1.8)

（b）$\mathbf{A} = \sum \mathbf{C}_i'(\mathbf{C}_i\mathbf{A}^{-1}\mathbf{C}_i')^{-1}\mathbf{C}_i$　(1c.1.9)

は同等であることがわかる．このとき，

$$(1c.1.9) \Rightarrow \mathbf{X'AX} = \sum \mathbf{Y}_i'(\mathbf{C}_i\mathbf{A}^{-1}\mathbf{C}_i')^{-1}\mathbf{Y}_i$$

となり，十分性が証明できる．

必要性を証明するためには，$(1c.1.7) \Rightarrow \mathbf{A} = \mathbf{C'BC}$ で，ここで，$\mathbf{B} = \mathrm{diag}(\mathbf{B}_1, \cdots, \mathbf{B}_k)$ は正則である．それゆえ，$\mathbf{B}^{-1} = \mathbf{CA}^{-1}\mathbf{C'}$ となり，これから，明らかに，$i \neq j$ にたいして $\mathbf{C}_i\mathbf{A}^{-1}\mathbf{C}_j' = 0$ で，かつ，$(\mathbf{C}_i\mathbf{A}^{-1}\mathbf{C}_i')^{-1} = \mathbf{B}_i$ であることが導かれる．

1c.2　固有方程式の根

$\mathbf{A}$ の固有方程式，あるいは，$\mathbf{A}$ の特性方程式とも呼ばれる $|\mathbf{A}-\lambda\mathbf{I}|=0$ を考える．$\mathbf{A}$ が $m\times m$ なら，この方程式は，λ の次数 $\leqslant m$ である．任意の根 λ_i（$\mathbf{A}$ の固有根，特性根，あるいは固有値と呼ばれる）に対応して，$\mathbf{AP}_i=\lambda_i\mathbf{P}_i$ で $\mathbf{P}_i^*\mathbf{P}_i=1$ なる非零の列ベクトル $\mathbf{P}_i$（固有ベクトル，あるいは特性ベクトルと呼ばれる）が存在する．$\mathbf{A}$ が実対称行列であるとき，次の諸結果は興味がある．

（i）$R(\mathbf{A})=r$ ならば，方程式 $|\mathbf{A}-\lambda\mathbf{I}|=0$ は，ゼロを $m-r$ 重根としてもつ．

（ii）すべての固有根は実数で，固有ベクトルも実数にとることができる．

i を虚数として，$\mathbf{A}(\mathbf{X}+i\mathbf{Y})=(\lambda+i\mu)(\mathbf{X}+i\mathbf{Y})$ とする．実数部分と虚数部分をそれぞれ等しいとおき，それぞれに $\mathbf{Y}'$ と $\mathbf{X}'$ を前からかけ，減算すると，$\mu=0$ を得る．$\mathbf{Y}$ はゼロにとることができる．

（iii）λ_i,λ_j $(\lambda_i\neq\lambda_j)$ に対応する固有ベクトル $\mathbf{P}_i$ と $\mathbf{P}_j$ は直交する．

定義から，

$$\mathbf{AP}_i=\lambda_i\mathbf{P}_i,\qquad \mathbf{AP}_j=\lambda_j\mathbf{P}_j$$

である．第1の式に $\mathbf{P}_j'$ を，第2の式に $\mathbf{P}_i'$ を前からかけて，減算をすると，$(\lambda_i-\lambda_j)\mathbf{P}_i'\mathbf{P}_j=0\Rightarrow\mathbf{P}_i'\mathbf{P}_j=0$ を得る．

（iv）$\mathbf{A}$ が p. d. ならば，すべての根は正である．$\mathbf{A}$ が n. n. d. ならば，すべての根は非負である．

$$\mathbf{AP}_i=\lambda_i\mathbf{P}_i\Rightarrow\mathbf{P}_i'\mathbf{AP}_i=\lambda_i\mathbf{P}_i'\mathbf{P}_i$$

を考える．$\mathbf{A}$ が p. d. なら，$\mathbf{P}_i'\mathbf{AP}_i$ と $\mathbf{P}_i'\mathbf{P}_i$ はともに正である．したがって，$\lambda_i>0$．$\mathbf{A}$ が n. n. d. なら $\mathbf{P}_i'\mathbf{AP}_i$ は非負である．したがって $\lambda_i\geqslant 0$．

（v）$\mathbf{X}$ を任意の非零ベクトルとするとき，線形部分区間 $\mathscr{M}(\mathbf{X},\mathbf{AX},\mathbf{A}^2\mathbf{X},\cdots)$ に属する固有ベクトル $\mathbf{Y}$ が存在する．

ベクトル $\mathbf{X},\mathbf{AX},\cdots$ は，すべてが独立であるというわけにはいかない．いま，k を，

$$\mathbf{A}^k\mathbf{X}+b_{k-1}\mathbf{A}^{k-1}\mathbf{X}+\cdots+b_0\mathbf{X}=0 \tag{1c.2.1}$$

が成り立つ最小の値とする．(1c.2.1) を因数分解すると，

$$(\mathbf{A}-\mu_1\mathbf{I})\cdots(\mathbf{A}-\mu_k\mathbf{I})\mathbf{X}=0\Rightarrow(\mathbf{A}-\mu_1\mathbf{I})\mathbf{Y}=0$$

ただし，

$$\mathbf{Y}=(\mathbf{A}-\mu_2\mathbf{I})\cdots(\mathbf{A}-\mu_k\mathbf{I})\mathbf{X}\neq 0 \tag{1c.2.2}$$

を得る．さらに，$(\mathbf{A}-\mu_1\mathbf{I})\mathbf{Y}=0$ ということは，$\mathbf{Y}$ が固有根 μ_1 に対応する固有ベクトルであることを意味する．$\mathbf{A}$ は実行列であるから，(ii) の結果を用いて，μ_1 は実数である．同様に，各 μ_i は実数となるから，方程式 (1c.2.2) は，

$$\mathbf{Y}\in\mathscr{M}(\mathbf{A}^{k-1}\mathbf{X},\cdots,\mathbf{X})$$

を示すことになる．

1c.3　行列の正準化

（i）実対称行列　$\mathbf{A}$ を $m\times m$ 実対称行列とする．このとき，Λ を対角行列とすると，$\mathbf{P}'\mathbf{AP}=\Lambda$ すなわち $\mathbf{A}=\mathbf{P}\Lambda\mathbf{P}'$ とするような直交行列 $\mathbf{P}$ が存在する．

$\lambda_1\geqslant\cdots\geqslant\lambda_m$ で，重根を含めた $\mathbf{A}$ の固有根を表すと，Λ の i 番目の対角要素は λ_i，$\mathbf{P}$ の i 番目の列ベクトルは λ_i に対応する固有ベクトル $\mathbf{P}_i$，とできることを実際に示そう．

いま，s 個の直交ベクトル $\mathbf{P}_1,\cdots,\mathbf{P}_s$ が存在し，

$$\mathbf{AP}_i=\lambda_i\mathbf{P}_i,\qquad i=1,\cdots,s \tag{1c.3.1}$$

を満足するとしよう．(1c.3.1) の結果から，$\mathbf{A}^2\mathbf{P}_i=\lambda_i\mathbf{AP}_i=\lambda_i^2\mathbf{P}_i,\cdots,\mathbf{A}^r\mathbf{P}_i=\lambda_i^r\mathbf{P}_i,\cdots$ を得る．$\mathscr{M}(\mathbf{P}_1,\cdots,\mathbf{P}_s)$ に $\perp^{\mathrm{er}}$ であるベクトル $\mathbf{X}$ を選ぶと，

$$\mathbf{X}'\mathbf{A}^r\mathbf{P}_i=\mathbf{X}'\lambda_i^r\mathbf{P}_i=\lambda_i^r\mathbf{X}'\mathbf{P}_i=0$$

が，すべての r と $i=1,\cdots,s$ にたいして成立する．したがって，

$$\mathscr{M}(\mathbf{X},\mathbf{AX},\mathbf{A}^2\mathbf{X},\cdots)\perp^{\mathrm{er}}\mathscr{M}(\mathbf{P}_1,\cdots,\mathbf{P}_s). \tag{1c.3.2}$$

[(v), 1c.2] の結果から，$\mathbf{P}_{s+1}\in\mathscr{M}(\mathbf{X},\mathbf{AX},\mathbf{A}^2\mathbf{X},\cdots)$ なる固有ベクトル $\mathbf{P}_{s+1}$ が存在し，(1c.3.2) を見れば，これは $\mathbf{P}_1,\cdots,\mathbf{P}_s$ に直交することがわかる．

$\mathbf{P}_1$ をそれから出発する任意の固有ベクトルにとることができるから，以上によって，

$$\mathbf{AP}_i=\lambda_i\mathbf{P}_i,\qquad i=1,\cdots,m \tag{1c.3.3}$$

を満足する，m 個の互いに直交する固有ベクトル $\mathbf{P}_1,\cdots,\mathbf{P}_m$ の存在を証明したことになる．また，上式は，$\mathbf{P}$ を $\mathbf{P}_i$ を列とする直交行列，Λ を λ_i を i 番目の対角要素とする対角行列とすると，次のように書ける．

$$\mathbf{AP}=\mathbf{P}\Lambda,\qquad \mathbf{PP}'=\mathbf{I}. \tag{1c.3.4}$$

(1c.3.4) の結果は，次の有用な形式にあらわすことができる．

$$\begin{aligned}\mathbf{A}&=\mathbf{P}\Lambda\mathbf{P}'=\lambda_1\mathbf{P}_1\mathbf{P}_1'+\cdots+\lambda_m\mathbf{P}_m\mathbf{P}_m'\\ \mathbf{I}&=\mathbf{PP}'\quad=\quad\mathbf{P}_1\mathbf{P}_1'+\cdots+\quad\mathbf{P}_m\mathbf{P}_m'\end{aligned} \tag{1c.3.5}$$

これは，$\mathbf{A}$ のスペクトル分解として知られている．(1c.3.4) の結果から，2次形式 $\mathbf{X}'\mathbf{AX}$ を

$$\mathbf{Y}'\Lambda\mathbf{Y}=\lambda_1y_1^2+\cdots+\lambda_my_m^2 \tag{1c.3.6}$$

に変換する直交変換 $\mathbf{Y}=\mathbf{P}'\mathbf{X}$ が存在することも言える．必ずしも直交ではない変換によって，2次形式はいつも2乗項だけを含む単純な形に標準化できることは，(1c.1.1) で証明した．

$\mathbf{P}'\mathbf{AP}=\Lambda$ であるから，$\mathbf{P}'(\mathbf{A}-\lambda\mathbf{I})\mathbf{P}=\Lambda-\lambda\mathbf{I}$ となり，これから，

$$|\mathbf{A}-\lambda\mathbf{I}|=|\Lambda-\lambda\mathbf{I}|=(\lambda_1-\lambda)\cdots(\lambda_m-\lambda) \tag{1c.3.7}$$

と書ける．すなわち，定数 $\lambda_1,\cdots,\lambda_m$ は重根を含んで $|\mathbf{A}-\lambda\mathbf{I}|=0$ の根と等しく，したがって，行列 Λ が何であるかがわかった．

もし λ_i が r 重根であるなら，方程式 $\mathbf{AX}=\lambda_i\mathbf{X}$ を満足する，r 個の直交ベクトルが存在するはずである．それ以上に1つ余計のベクトルがあるなら，[(iii), 1c. 2] によって，それはすべての $\mathbf{P}_i$ に直交することになるから，零でなければならない．したがって，$\mathbf{A}-\lambda_i\mathbf{I}$ の階数は $m-r$ である．r 重の固有根 λ_i に対応する r 個の固有ベクトルは，固有部分空間とよばれる部分空間を生成する．2つの異なる根に対応する固有部分空間は直交するから，$\mathbf{A}$ は，ベクトル空間全体 E_m を，直交部分空間の直和

$$E_m = E_{r_1}\oplus\cdots\oplus E_{r_k}$$

に分解する．ここで，E_{r_i} は，r_i 重の固有値に対応する固有部分空間である．$r_1+\cdots+r_k=m$ であることに注意せよ．

（ii）　**実対称行列の対**　$\mathbf{A}$ と $\mathbf{B}$ は $m\times m$ 実対称行列とし，そのうち $\mathbf{B}$ は p. d. とする．このとき，Λ を対角行列とすると，$\mathbf{A}=\mathbf{R}^{-1\prime}\Lambda\mathbf{R}^{-1}$，$\mathbf{B}=\mathbf{R}^{-1\prime}\mathbf{R}^{-1}$ とするような行列 $\mathbf{R}$ が存在する．

$\lambda_1\geqslant\cdots\geqslant\lambda_m$ を $|\mathbf{A}-\lambda\mathbf{B}|=0$ の根とする．このとき，Λ の i 番目の対角要素は λ_i で，$\mathbf{R}$ の i 番目の列ベクトル $\mathbf{R}_i$ は方程式 $\mathbf{AR}_i=\lambda_i\mathbf{BR}_i$ を満足することが証明できる．

$\mathbf{B}$ は p. d. であるから，$\mathbf{B}=\mathbf{T}'\mathbf{T}$ とするような正則行列 $\mathbf{T}$ が存在する．

$$|\mathbf{A}-\lambda\mathbf{B}|=0=|\mathbf{A}-\lambda\mathbf{T}'\mathbf{T}|\Leftrightarrow|\mathbf{T}^{-1\prime}\mathbf{AT}^{-1}-\lambda\mathbf{I}|=0$$

を考える．$\mathbf{C}=\mathbf{T}^{-1\prime}\mathbf{AT}^{-1}$ とおくと，これは対称である．したがって，(i) を用いると，行列 $\mathbf{P}$ が存在して，

$$\mathbf{P}'\mathbf{CP}=\Lambda,\qquad \mathbf{P}'\mathbf{P}=\mathbf{I}$$

$$\mathbf{P}'\mathbf{T}'^{-1}\mathbf{AT}^{-1}\mathbf{P}=\Lambda,\qquad \mathbf{P}'\mathbf{P}=\mathbf{I}\tag{1c. 3. 8}$$

とできる．$\mathbf{R}=\mathbf{T}^{-1}\mathbf{P}$ にとると，(1c. 3. 8) は

$$\mathbf{R}'\mathbf{AR}=\Lambda,\qquad \mathbf{R}'\mathbf{BR}=\mathbf{I}\tag{1c. 3. 9}$$

と書け，これは求める結果を与える．

$\mathbf{R}^{-1\prime}$ の i 番目の列を $\mathbf{S}_i$ と書くと，(1c. 3. 5) と同様に，(1c. 3. 9) の別の表現として，

$$\mathbf{A}=\lambda_1\mathbf{S}_1\mathbf{S}_1'+\cdots+\lambda_m\mathbf{S}_m\mathbf{S}_m'$$

$$\mathbf{B}=\mathbf{S}_1\mathbf{S}_1'+\cdots+\mathbf{S}_m\mathbf{S}_m',\tag{1c. 3. 10}$$

を得る．

(1c. 3. 10) の結果からまた，変換 $\mathbf{Y}=\mathbf{R}^{-1}\mathbf{X}$ が存在して，2つの2次形式 $\mathbf{X}'\mathbf{AX}$ と $\mathbf{X}'\mathbf{BX}$ が同時に，2乗項だけを含むより単純な形式に標準化できることが導ける．すなわち

$$\mathbf{X}'\mathbf{AX}\to\lambda_1y_1^2+\cdots+\lambda_my_m^2$$

$$\mathbf{X}'\mathbf{BX}\to y_1^2+\cdots+y_m^2.\tag{1c. 3. 12}$$

交換可能行列　実2次形式は，直交変換によって，より単純な形式に標準化できることを (i) で示した．次に当然おこる疑問は，2つの2次形式が，1つの直交変換によって，同時により単純な形式に標準化できるか否かということである．その答は，交換可能行列 $\mathbf{A}$ と $\mathbf{B}$ については，すなわち，$\mathbf{AB}=\mathbf{BA}$ なら，イエスということである．

（iii）　**交換可能行列**　$\mathbf{A}$ と $\mathbf{B}$ を対称行列とする．$\mathbf{C}$ を直交行列とするとき，$\mathbf{C}'$ と $\mathbf{C}$ を前と後からかけることによって，$\mathbf{A}$ と $\mathbf{B}$ を同時に対角化できるための必要十分条件は，$\mathbf{AB}=\mathbf{BA}$（すなわち，$\mathbf{A}$ と $\mathbf{B}$ は交換可能）ということである．

$\mathbf{Q}_1$ を $\mathbf{B}$ の固有ベクトルとし，$\mathbf{AB}=\mathbf{BA}$ とする．すると，ある λ について，$\mathbf{BQ}_1=\lambda\mathbf{Q}_1$ を得る．$\mathbf{A}$ をかけると，

$$\mathbf{ABQ}_1=\lambda\mathbf{AQ}_1=\mathbf{BAQ}_1$$

となるから，$\mathbf{AQ}_1$，同時に，$\mathbf{A}^2\mathbf{Q}_1,\mathbf{A}^3\mathbf{Q}_1,\cdots$ は，同じ固有値 λ に対応する $\mathbf{B}$ の固有空間に属する．しかし，[(v), 1c. 2] によって，$\mathscr{M}(\mathbf{Q}_1,\mathbf{AQ}_1,\cdots)$ は $\mathbf{A}$ の固有ベクトルを含む．したがって，$\mathbf{A}$ と $\mathbf{B}$ には共通の固有ベクトルがあり，それを $\mathbf{P}_1$ であらわす．いま，$\mathbf{B}$ の固有ベクトルとして，$\mathbf{P}_1$ に $\perp^{er}$ な $\mathbf{Q}_2$ を選ぶ．同じ議論を適用すると，$\mathbf{P}_1$ に $\perp^{er}$ な共通の固有ベクトル $\mathbf{P}_2$ を得，次々に進む．$\mathbf{P}_1,\mathbf{P}_2,\cdots$ を列とする行列は，正確に，求める行列 $\mathbf{C}$ であり，これによって十分性が証明された．

$\mathbf{M}_1$ と $\mathbf{M}_2$ を対角行列とし，$\mathbf{C}'\mathbf{AC}=\mathbf{M}_1$，$\mathbf{C}'\mathbf{BC}=\mathbf{M}_2$ ならば，$\mathbf{AB}=(\mathbf{CM}_1\mathbf{C}')(\mathbf{CM}_2\mathbf{C}')=\mathbf{CM}_1\mathbf{M}_2\mathbf{C}'=\mathbf{CM}_2\mathbf{M}_1\mathbf{C}'=\mathbf{CM}_2\mathbf{C}'\mathbf{CM}_1\mathbf{C}'=\mathbf{BA}$ を得，必要性が証明される．

エルミート行列　$\mathbf{A}^*$ を $\mathbf{A}$ の共役転置とするとき，$\mathbf{A}=\mathbf{A}^*$ なら行列 $\mathbf{A}$ はエルミート行列と言われる．エルミート行列は，$\mathbf{A}$ が複素要素をもつときの対称行列とも考えられる．エルミート行列についての次の諸結果は (i), (ii) とまったく同様に証明できる．

（iv）　$\mathbf{A}$ を次数 m のエルミート行列とする．このとき

（a）　$\mathbf{A}$ の固有値は実数である．

（b）　ユニタリ行列 $\mathbf{U}$（すなわち，$\mathbf{U}^*\mathbf{U}=\mathbf{I}$）が存在して，$\mathbf{A}=\mathbf{U}\Lambda\mathbf{U}^*$ と書ける．ここで，Λ は $\mathbf{A}$ の固有値を要素とする対角行列である．

（c）　$\mathbf{B}$ をエルミート行列で，p. d.（すなわち，非零であるすべての $\mathbf{X}$ にたいして $\mathbf{X}^*\mathbf{BX}>0$）とする．このとき，正則行列 $\mathbf{P}$ が存在して，$\mathbf{P}^*\mathbf{BP}=\mathbf{I}$，$\mathbf{P}^*\mathbf{AP}=\Lambda$ とできる．ここで，Λ は行列 $\mathbf{B}^{-1}\mathbf{A}$ の固有値を要素とする対角行列である．

（v）　**特異値分解**　$\mathbf{A}$ を階数 r の $m\times n$ 行列とする．このとき，$\mathbf{A}$ は，

$$\mathbf{A}=\mathbf{U}\Lambda\mathbf{V}^*\tag{1c. 3. 11}$$

とあらわされる．ここで，Λ は正の対角要素をもつ $r\times r$ の対角行列，$\mathbf{U}$ は $\mathbf{U}^*\mathbf{U}=\mathbf{I}$ であるような $m\times r$ 行列，$\mathbf{V}$ は $\mathbf{V}^*\mathbf{V}=\mathbf{I}$ であるような $n\times r$ 行列である．

$\mathbf{U}_i$，$i=1,\cdots,r$ を，$\mathbf{AA}^*$ のゼロでない固有値 λ_i^2，$i=1,\cdots,r$ に対応する正規直交系の固有ベクトルとする．さらに，$\mathbf{V}_i=\lambda_i^{-1}\mathbf{A}^*\mathbf{U}_i$ とおく．このとき，$\mathbf{V}_i$，$i=1,\cdots,r$ も正規直交系となる．

$\mathbf{U}_{r+1},\cdots,\mathbf{U}_m$ を，$\mathbf{U}_1,\cdots,\mathbf{U}_m$ が正規直交系ベクトルの完全集合になるようにつけ加える，すなわち，$\mathbf{U}_1\mathbf{U}_1^*+\cdots+\mathbf{U}_m\mathbf{U}_m^*=\mathbf{I}$ とする．このとき，$\mathbf{U}=(\mathbf{U}_1\,\vdots\cdots\vdots\,\mathbf{U}_r)$，$\mathbf{V}=(\mathbf{V}_1\,\vdots\cdots\vdots\,\mathbf{V}_r)$ とすると，

$$\begin{aligned}\mathbf{A}&=(\mathbf{U}_1\mathbf{U}_1^*+\cdots+\mathbf{U}_m\mathbf{U}_m^*)\mathbf{A}\\&=(\mathbf{U}_1\mathbf{U}_1^*+\cdots+\mathbf{U}_r\mathbf{U}_r^*)\mathbf{A},\qquad i>r\text{ のとき }\mathbf{U}_i^*\mathbf{A}=\mathbf{0}\text{ となるから，}\\&=\lambda_1\mathbf{U}_1\mathbf{V}_1^*+\cdots+\lambda_r\mathbf{U}_r\mathbf{V}_r^*=\mathbf{U}\Lambda\mathbf{V}^*\end{aligned}$$

を得る．Λ の対角線上の値は，$\mathbf{A}$ の特異値と呼ばれる．$\mathbf{V}_i$，$i=1,\cdots,r$ は，$\mathbf{A}^*\mathbf{A}$ の固有値 λ_i^2，$i=1,\cdots,r$ に対応する正規直交系の固有ベクトルであることは容易にわかる．$\mathbf{A}$を表現する次の別の方法に注目せよ．

$$\begin{aligned}\mathbf{A}&=\mathbf{U}\Lambda\mathbf{V}^*=\lambda_1\mathbf{U}_1\mathbf{V}_1^*+\cdots+\lambda_r\mathbf{U}_r\mathbf{V}_r^*\\&=\mathbf{P}\begin{pmatrix}\Lambda&0\\0&0\end{pmatrix}\mathbf{Q}^*,\qquad\mathbf{P}\text{ と }\mathbf{Q}\text{ はユニタリ行列.}\end{aligned}\tag{1c. 3. 12}$$

（vi）基本分解定理　<u>$\mathbf{A}$ を次数 n の正方行列とする．このとき，$\mathbf{P}^*\mathbf{A}\mathbf{P}$ を三角行列とするようなユニタリ行列 $\mathbf{P}$ が存在する．</u>

いま次数 $n-1$ の行列にたいしてこの命題は正しいとする．次数 n の行列についても正しいことを証明しよう．

λを1つの固有値，$\mathbf{x}$ をそれに対応する固有ベクトル，$\mathbf{Q}$ を $\mathbf{Qx}=k\mathbf{e}_1$ であるようなユニタリ行列とする．ただし，$\mathbf{e}_1'=(1,0,\cdots,0)$ で k は零でない定数である．このとき，

$$\begin{aligned}\mathbf{Ax}=\lambda\mathbf{x}&\Leftrightarrow\mathbf{QAx}=\lambda\mathbf{Qx}\Leftrightarrow\mathbf{QAQ}^*\mathbf{Qx}=\lambda\mathbf{Qx},\\&\Leftrightarrow\mathbf{QAQ}^*k\mathbf{e}_1=\lambda k\mathbf{e}_1\Leftrightarrow\mathbf{QAQ}^*\mathbf{e}_1=\lambda\mathbf{e}_1.\end{aligned}$$

したがって，$\mathbf{QAQ}^*$ は次の形に書ける．

$$\left(\begin{array}{c|c}\lambda&\mathbf{b}^*\\\hline 0&\mathbf{A}_{n-1}\end{array}\right)$$

仮説により，$\mathbf{R}^*\mathbf{A}_{n-1}\mathbf{R}$ を三角行列とするような，次数 $n-1$ のユニタリ行列 $\mathbf{R}$ は存在する．それゆえ，求める行列 $\mathbf{P}$ として

$$\mathbf{Q}^*\begin{pmatrix}1&0'\\0&\mathbf{R}\end{pmatrix}$$

をとれば，命題は証明される．

（vii）非対称行列　固有方程式 $|\mathbf{A}-\lambda\mathbf{I}|=0$ は m 個の根をもち，そのうちいくつかは，$\mathbf{A}$ が実行列であっても複素数になることがある．根 λ_i に対応して2つのベクトル

$$\mathbf{AP}_i=\lambda_i\mathbf{P}_i,\qquad\mathbf{A}'\mathbf{Q}_i=\lambda_i\mathbf{Q}_i\tag{1c. 3. 13}$$

があり，それぞれ，右および左固有ベクトルと呼ばれる．すべての根は異なると仮定し，

$$\mathbf{P}=(\mathbf{P}_1\,\vdots\cdots\vdots\,\mathbf{P}_m),\qquad\mathbf{Q}=(\mathbf{Q}_1\,\vdots\cdots\vdots\,\mathbf{Q}_m)$$

と書くと，$\mathbf{AP}=\mathbf{P}\Lambda$，$\mathbf{A}'\mathbf{Q}=\mathbf{Q}\Lambda$ が導かれる．

（a）<u>$\mathbf{P}_i$ はすべて独立で，$\mathbf{Q}_i$ もまたそうである．</u>いま，$c_1\mathbf{P}_1+\cdots+c_m\mathbf{P}_m=0$ を満足するような定数が存在するとする．(1c. 3. 13) を用いると，$c_1\lambda_1\mathbf{P}_1+\cdots+c_m\lambda_m\mathbf{P}_m=0$ を得る．この項の1つ，たとえば，m 番目を，$\mathbf{P}_i$ を結合する上の2つの1次結合から除去すると，$d_1\mathbf{P}_1+\cdots+d_{m-1}\mathbf{P}_{m-1}=0\Rightarrow d_1\lambda_1\mathbf{P}_1+\cdots+d_{m-1}\lambda_{m-1}\mathbf{P}_{m-1}=0$ を得る．さらにもう1つの項を除くと，ついには，$\mathbf{P}_i$ の1つが零になることを示すことができ，これは矛盾である．したがって，$\mathbf{P}$ は正則で，$\mathbf{Q}$ もまた正則になり，次の表現に到達する．

$$\mathbf{A}=\mathbf{P}\Lambda\mathbf{P}^{-1},\qquad\mathbf{A}'=\mathbf{Q}\Lambda\mathbf{Q}^{-1}.$$

（b）$i\neq j$ にたいして $\mathbf{P}_i'\mathbf{Q}_j=0$ で，かつ，$\mathbf{P}_i'\mathbf{Q}_i\neq0$.

$$\mathbf{AP}_i=\lambda_i\mathbf{P}_i\Rightarrow\mathbf{Q}_j'\mathbf{AP}_i=\lambda_i\mathbf{Q}_j'\mathbf{P}_i=\lambda_i\mathbf{P}_i'\mathbf{Q}_j$$

$$\mathbf{A}'\mathbf{Q}_j=\lambda_j\mathbf{Q}_j\Rightarrow\mathbf{Q}_j'\mathbf{AP}_i=\lambda_j\mathbf{Q}_j'\mathbf{P}_i=\lambda_j\mathbf{P}_i'\mathbf{Q}_j$$

これから，$\lambda_i\neq\lambda_j$ であるから，$\mathbf{P}_i'\mathbf{Q}_j=0$ を得る．さらに，すべての j にたいして $\mathbf{P}_i'\mathbf{Q}_j=0$ なら，$\mathbf{P}_i=\mathbf{0}$ となり，これは正しくない．よって $\mathbf{Q}_i'\mathbf{P}_i\neq0$.

（c）<u>$\mathbf{P}_i'\mathbf{Q}_i=d_i$，$\mathbf{D}$ を d_i を要素とする対角行列，Δ を $\delta_i=\lambda_i/d_i$ を要素とする対角行列とする．このとき，$\mathbf{A}$ のスペクトル分解は次のようになる．</u>

$$\mathbf{A}=\delta_1\mathbf{P}_1\mathbf{Q}_1'+\cdots+\delta_m\mathbf{P}_m\mathbf{Q}_m'\tag{1c. 3. 14}$$

$\mathbf{Q}'\mathbf{P}=\mathbf{D}$ であるから，$\mathbf{P}^{-1}(\mathbf{Q}')^{-1}=\mathbf{D}^{-1}$，よって，$\mathbf{P}^{-1}=\mathbf{D}^{-1}\mathbf{Q}'$ となる．これを $\mathbf{A}=\mathbf{P}\Lambda\mathbf{P}^{-1}$ における $\mathbf{P}^{-1}$ に代入すると，(1c. 3. 5) と同様に，

$$\mathbf{A}=\mathbf{P}\Lambda\mathbf{D}^{-1}\mathbf{Q}'=\mathbf{P}\Delta\mathbf{Q}'=\Sigma\delta_i\mathbf{P}_i\mathbf{Q}_i'$$

を得る．非対称行列 $\mathbf{A}$ のこの分解は，固有根のすべてが異ならないときには，それほど簡単にはならない．この問題を研究するには，Aitken (1944)，Gantmacher (1959)，Turnbull and Aitken (1932) の本を参照されたい．

(1c. 3. 5)，(1c. 3. 14) の結果の1つの応用は $\mathbf{A}$ のベキの簡単な表現を得ることにある．(1c. 3. 14) を考慮して，$\mathbf{A}$ をかけると

$$\mathbf{A}^2=\Sigma\frac{\lambda_i}{d_i}\mathbf{P}_i\mathbf{Q}_i'\mathbf{A}=\Sigma\frac{\lambda_i^2}{d_i}\mathbf{P}_i\mathbf{Q}_i'\tag{1c. 3. 15}$$

を得る．同様にして，

$$\mathbf{A}^n=\Sigma\frac{\lambda_i^n}{d_i}\mathbf{P}_i\mathbf{Q}_i'.$$

対称行列については，(1c. 3. 5) を用いると，

$$\mathbf{A}^n=\Sigma\lambda_i^n\mathbf{P}_i\mathbf{P}_i'$$

を得る．(1c. 3. 15) の遺伝学への応用は Fisher (1949) の本に与えられている．

（viii）**ケーレイ＝ハミルトンの定理**　どの正方行列 **A** も，それ自身の特性方程式を満足する．

恒等式

$$(\mathbf{A}-\lambda\mathbf{I})\mathrm{adj}(\mathbf{A}-\lambda\mathbf{I})=|\mathbf{A}-\lambda\mathbf{I}|\mathbf{I} \tag{1c.3.16}$$

を考える（adj**B** とは **B** の余因子の行列である）．定義により，

$$\mathrm{adj}(\mathbf{A}-\lambda\mathbf{I})=\mathbf{C}_0+\lambda\mathbf{C}_1+\cdots+\lambda^{n-1}\mathbf{C}_{n-1},$$

ここで，$\mathbf{C}_i$ は適当に選ばれた行列で $|\mathbf{A}-\lambda\mathbf{I}|=a_0+a_1\lambda+\cdots+a_n\lambda^n$ は特性多項式である．(1c.3.16) の λ の係数を比較して，

$$\mathbf{AC}_0=a_0\mathbf{I},\quad \mathbf{AC}_1-\mathbf{C}_0=a_1\mathbf{I},\quad \cdots,\quad -\mathbf{C}_{n-1}=a_n\mathbf{I}$$

を得る．第1式に **I** を，第2式に **A** を，第3式に $\mathbf{A}^2$ を…，というようにかけて加えると

$$a_0\mathbf{I}+a_1\mathbf{A}+\cdots+a_n\mathbf{A}^n=0 \tag{1c.3.17}$$

を得，これで求める結果が証明された．

$\mathbf{A}^{-1}$ が存在するなら，(1c.3.17) に $\mathbf{A}^{-1}$ をかけると，

$$-a_0\mathbf{A}^{-1}=a_1\mathbf{I}+a_2\mathbf{A}+\cdots+a_n\mathbf{A}^{n-1} \tag{1c.3.18}$$

が得られ，これは，$\mathbf{A}^{-1}$ を，**A** の多項式に展開する表現になる．

（ix）**正規行列の分解**　**A** を $n\times n$ の正規行列，つまり $\mathbf{AA}^*=\mathbf{A}^*\mathbf{A}$ とする．このとき Λ を対角行列として，

$$\mathbf{A}=\mathbf{U}\Lambda\mathbf{U}^*$$

とするようなユニタリ行列 **U** が存在する．

［(vi), **1c.3**］の基本分解定理を用いる．**P** をユニタリ，**T** を三角行列とすると，

$$\mathbf{A}=\mathbf{PTP}^*. \tag{1c.3.12b}$$

A は正規であるから，

$$\mathbf{AA}^*=\mathbf{A}^*\mathbf{A}\Leftrightarrow\mathbf{T}^*\mathbf{T}=\mathbf{TT}^*$$

となり，**T** は対角行列となることがわかるから，命題は証明された．

エルミート行列の分解についての［(iv), **1c.3**］の結果は，［(vi), **1c.3**］の直接の帰結でもあることに注意せよ．(1c.3.12b) からは，**A** がエルミート行列で **T** が対角形になるなら，$\mathbf{T}=\mathbf{T}^*$ を得る．［(i), **1c.3**］はこの直接の系である．

（x）**ジョルダン分解**　**A** を次数 m の正方行列とする．このとき，

$$\mathbf{T}^{-1}\mathbf{AT}=\mathrm{diag}(\Lambda_1,\Lambda_2,\cdots,\Lambda_k)$$

とするような正則行列 **T** が存在する．ここで，

$$\Lambda_i=\begin{pmatrix}\lambda_i & 1 & 0 & \cdots & 0\\ 0 & \lambda_i & 1 & \cdots & 0\\ & \cdots & \cdots & \cdots & \\ 0 & 0 & \cdots & \cdots & \lambda_i\end{pmatrix}$$

で，λ_i は **A** の固有値である．

証明は，Gantmacher (1959) を見よ．

非負行列（スペクトル定理）　$m\times n$ 行列 $\mathbf{A}=(a_{ij})$ は，すべての i と j について，$a_{ij}\geqq 0$ なら非負，$a_{ij}>0$ なら正と呼ばれる．

非負行列は $\mathbf{A}\geqq 0$ で，正行列は $\mathbf{A}>0$ であらわされる．$\mathbf{C}-\mathbf{D}$ が非負のときは $\mathbf{C}\geqq\mathbf{D}$，$\mathbf{C}-\mathbf{D}$ が正のときは $\mathbf{C}>\mathbf{D}$ という記号を用いる．

行列 **A** は列と行に同じ順列置換を施して

$$\begin{pmatrix}\mathbf{B} & 0\\ \mathbf{C} & \mathbf{D}\end{pmatrix}$$

と書くことができれば，可約と言われる．ここで，**0** は零行列，**B**, **D** は正方行列である．可約でなければ，**A** は既約である．

（xi）**ペロンの定理**　正の正方行列 **A** には，特性方程式の単純根で他のすべての固有値の絶対値よりも大きい，正の固有値 $\lambda_0(\mathbf{A})$ がある．この“最大の”根 $\lambda_0(\mathbf{A})$ には，正の（すなわち，すべての座標が正の）固有ベクトルが対応する．

（xii）**フロベニウスの定理**　既約行列 $\mathbf{A}\geqq 0$ には，特性方程式の単純根で他の固有値の絶対値よりは小さくない，正の固有値 $\lambda_0(\mathbf{A})$ がつねに存在する．この“最大”根 $\lambda_0(\mathbf{A})$ には，正の固有ベクトルが対応する．

（xiii）既約行列 $\mathbf{A}\geqq 0$ は，2つの線形独立な非負の固有ベクトルをもつことはできない．

（xiv）非負行列 **A** は，非負の固有値 $\lambda_0(\mathbf{A})$ をつねにもち，**A** のすべての固有値はその絶対値が $\lambda_0(\mathbf{A})$ を超えない．この“最大”根には，非負の固有ベクトルが対応する．

（xv）$\mathbf{A}\geqq\mathbf{B}\geqq 0$ なら，$\lambda_0(\mathbf{A})\geqq\lambda_0(\mathbf{B})$ である．

（xvi）［**A** は既約で］$\mathbf{A}\geqq 0$，$\mu_0=\inf\{\mu:$ ある $\mathbf{f}>0$ にたいして $\mathbf{Af}\leqq\mu\mathbf{f}\}$ とすると，

$$\sup_{\substack{\mathbf{X}>0\\ \sum x_i=1}}\ \inf_{\substack{\mathbf{f}>0\\ \sum f_i=1}}\frac{\mathbf{f}'\mathbf{AX}}{\mathbf{f}'\mathbf{X}}=\inf_{\substack{\mathbf{f}>0\\ \sum f_i=1}}\ \sup_{\substack{\mathbf{X}>0\\ \sum x_i=1}}\frac{\mathbf{f}'\mathbf{AX}}{\mathbf{f}'\mathbf{X}}\qquad[=\mu_0=\lambda_0(\mathbf{A})]$$

が成り立つ．

メツラー行列　行列 $\mathbf{C}=(c_{ij})$ は，$c_{ij}\geqq 0$ $(i\neq j)$ のとき，メツラー行列とよばれる．これに $\beta\mathbf{I}$ を加え，β を十分大きくとると，

$$\mathbf{A}=\mathbf{C}+\beta\mathbf{I}\geqq 0$$

とできることは，容易にわかる．それゆえ，メツラー行列，すなわち，M 行列の研究は，非負行列のそれと同等になる．したがって，次を得る．

（xvii）（a）M 行列 **M** の固有値で最大の実数部分をもつものは，実数で，これには非負の固有ベクトルが対応する．

（b）$\sigma(\mathbf{M})$ を $\mathbf{M}$ の固有値の実数部分の最大値とする．このとき，$\sigma(\mathbf{M})<0$ になるのは，$\mathbf{MX}<0$ となるような $\mathbf{X}>0$ が存在するときだけである．

この命題の証明は Gantmacher (1959) を見よ．

1c.4 射影作用素

定義　$\mathscr{S}_1$ と $\mathscr{S}_2$ は互いに素な部分空間で $\mathscr{V}=\mathscr{S}_1\oplus\mathscr{S}_2$ とする．よって，任意のベクトル $\mathbf{X}\in\mathscr{V}$ は，$\mathbf{X}=\mathbf{X}_1+\mathbf{X}_2$，$\mathbf{X}_1\in\mathscr{S}_1$，$\mathbf{X}_2\in\mathscr{S}_2$ に一意的に分解されるとする．このとき，写像 $\mathbf{P}$，$\mathbf{X}\to\mathbf{X}_1$ $(\mathbf{PX}=\mathbf{X}_1)$ を，$\mathscr{S}_1$ の上への $\mathscr{S}_2$ に沿っての $\mathbf{X}$ の射影とよび，$\mathbf{P}$ を射影作用素とよぶ．

射影作用素 $\mathbf{P}$ に関して，次の諸命題を得る．

（i）$\mathbf{P}$ は線形変換で，かつ，一意的である．

明らかに，$\mathbf{P}(a\mathbf{X}+b\mathbf{Y})=a\mathbf{PX}+b\mathbf{PY}$ であり，これは，$\mathbf{P}$ が，同じ記号 $\mathbf{P}$ であらわされる行列によって表現できることと同等である．よって，$\mathbf{P}$ は必ず一意である．

（ii）$\mathbf{P}$ はベキ等行列である．すなわち，$\mathbf{P}^2=\mathbf{P}$．

任意のベクトル $\mathbf{X}$ にたいして

$$\mathbf{P}^2\mathbf{X}=\mathbf{P}(\mathbf{PX})=\mathbf{PX}_1=\mathbf{X}_1=\mathbf{PX}$$

となり，これから $\mathbf{P}^2=\mathbf{P}$ が導かれる．

（iii）$\mathbf{I}-\mathbf{P}$ は，$\mathscr{S}_2$ の上への $\mathscr{S}_1$ に沿っての射影子である．

（iv）$\mathscr{V}=\mathscr{S}_1\oplus\mathscr{S}_2\oplus\cdots\oplus\mathscr{S}_k$ とする．このとき，行列 $\mathbf{P}_1,\cdots,\mathbf{P}_k$ が存在して，(a) $\mathbf{P}_i^2=\mathbf{P}_i$，(b) $\mathbf{P}_i\mathbf{P}_j=0$，$i\neq j$，(c) $\mathbf{I}=\mathbf{P}_1+\cdots+\mathbf{P}_k$ とできる．

この結果は，一意的分解，

$$\mathbf{X}=\mathbf{X}_1+\cdots+\mathbf{X}_k\qquad \mathbf{X}_i\in\mathscr{S}_i,\qquad i=1,\cdots,k$$

から出る．

定義　$\mathscr{S}$ を $\mathscr{V}$ の部分空間，$\mathscr{S}^\perp$ を $\mathscr{S}$ の直交補空間とする．このとき，$\mathscr{V}=\mathscr{S}\oplus\mathscr{S}^\perp$ となる．$\mathscr{S}$ の上への $\mathscr{S}^\perp$ に沿っての射影作用素 $\mathbf{P}$ を直交射影子とよぶ．

直交射影に関しては，次の諸命題を得る．

（v）Σ を p.d. 行列として，$\mathscr{V}$ における内積を $(\mathbf{X},\mathbf{Y})=\mathbf{Y}^*\Sigma\mathbf{X}$ で定義する．このとき，$\mathbf{P}$ が直交射影子であるのは，

$$\text{(a)}\ \mathbf{P}^2=\mathbf{P},\qquad \text{かつ}\qquad \text{(b)}\ \Sigma\mathbf{P}\ \text{がエルミート}\qquad (1c.4.1)$$

のときだけである．

結果 (a) は，命題 (ii) から正しい．さらに，任意の $\mathbf{X},\mathbf{Y}$ にたいして，$\mathbf{PX}\in\mathscr{S}$ で，かつ $(\mathbf{I}-\mathbf{P})\mathbf{Y}\in\mathscr{S}^\perp$ である．すなわち，$\mathbf{X}^*\mathbf{P}^*\Sigma(\mathbf{I}-\mathbf{P})\mathbf{Y}=0\Leftrightarrow\mathbf{P}^*\Sigma(\mathbf{I}-\mathbf{P})=0\Leftrightarrow\mathbf{P}^*\Sigma\mathbf{P}=\Sigma\mathbf{P}$ は (b) を証明することになる．(1c.4.1) が成立するなら，$\mathbf{P}$ は $\mathscr{M}(\mathbf{P})$ の上へ直交補空間 $\mathscr{M}(\mathbf{I}-\mathbf{P})$ に沿ってベクトルを射影するものであることがすぐわかる．

注意　$\Sigma=\mathbf{I}$ なら，$\mathbf{P}$ が直交射影子になるのは，$\mathbf{P}$ がベキ等でエルミート行列のときのみである．

（vi）**直交射影子の陽表現**　行列 $\mathbf{A}$ の列によって張られる部分空間 $\mathscr{M}(\mathbf{A})$ を考え，その内積を命題 (v) のとおりとする．このとき，$\mathscr{M}(\mathbf{A})$ の上への直交射影子は，

$$\mathbf{P}=\mathbf{A}(\mathbf{A}^*\Sigma\mathbf{A})^-\mathbf{A}^*\Sigma\qquad (1c.4.2)$$

となり，(1c.4.2) に含まれる一般逆行列の選択のいかんにかかわらず，一意にきまる．

$\mathbf{X}\in\mathscr{V}$ とし，$\mathbf{X}=\mathbf{A}\alpha+\mathbf{A}^\perp\beta$ と書く．両辺に $\mathbf{A}^*\Sigma$ をかけると，

$$\mathbf{A}^*\Sigma\mathbf{A}\alpha=\mathbf{A}^*\Sigma\mathbf{X}$$

を得る．解として，$\alpha=(\mathbf{A}^*\Sigma\mathbf{A})^-\mathbf{A}^*\Sigma\mathbf{X}$ を選ぶと，$\mathbf{A}\alpha=\mathbf{A}(\mathbf{A}^*\Sigma\mathbf{A})^-\mathbf{A}^*\Sigma\mathbf{X}\Rightarrow\mathbf{P}=\mathbf{A}(\mathbf{A}^*\Sigma\mathbf{A})^-\mathbf{A}^*\Sigma$ が求める作用素となる．$\Sigma=\mathbf{I}$ のときは，

$$\mathbf{P}=\mathbf{A}(\mathbf{A}^*\mathbf{A})^-\mathbf{A}^*\qquad (1c.4.3)$$

となる．

$\mathbf{A}(\mathbf{A}^*\mathbf{A})^-\mathbf{A}^*$ は，その一般逆行列の選択のいかんにかかわらず，対称でエルミート行列で，かつ，不変であることが，すでに [(viii), 1b.5] に示されている．

（v）$\mathbf{P}_1$ と $\mathbf{P}_2$ を射影子とする．このとき，$\mathbf{P}_1+\mathbf{P}_2$ が射影子となるのは，$\mathbf{P}_1\mathbf{P}_2=\mathbf{P}_2\mathbf{P}_1=0$ のときのみである．

1c.5 一般逆行列に関するその他の結果

要素が任意の体に属する行列 $\mathbf{A}$ の一般逆行列の普遍的な定義は 1b.5 に与えた．$\mathbf{A}^-$ であらわされるそのような逆行列はつねに存在するが，一意でないことが示された．本節では，ある種の最適化問題を解くのに有用な，一般逆行列の部分集合を定義する．この目的のためには，複素数を要素とする行列 $\mathbf{A}$ を考える．

（i）**ノルム最小型一般逆行列**　$\mathbf{AX}=\mathbf{Y}$ は解をもつ方程式で，$\mathbf{G}$ は $\mathbf{A}$ の一般逆行列，$\mathbf{GY}$ は $\mathbf{AX}=\mathbf{Y}$ の解のなかでノルム最小のものとする．ベクトル α のノルムを，p.d. 行列 $\mathbf{N}$ を用いて，$\|\alpha\|=(\alpha^*\mathbf{N}\alpha)^{\frac{1}{2}}$ で定義すると，$\mathbf{G}$ は条件

$$\mathbf{AGA}=\mathbf{A},\qquad (\mathbf{GA})^*\mathbf{N}=\mathbf{NGA}\qquad (1c.5.1)$$

を満足する．このような $\mathbf{G}$ を $\mathbf{A}_{\bar{m}(\mathbf{N})}$ であらわし，$\mathbf{N}=\mathbf{I}$ のとき $\mathbf{A}_{\bar{m}}$ と書く．

[(ii), 1b.5] を用いると，$\mathbf{AX}=\mathbf{Y}$ の一般解は，$\mathbf{GY}+(\mathbf{I}-\mathbf{GA})\mathbf{Z}$ となる．ここで，$\mathbf{Z}$ は任意のベクトルである．$\mathbf{GY}$ のノルムが最小であるときは，

$$\begin{aligned}\|\mathbf{GY}\|\leqslant\|\mathbf{GY}+(\mathbf{I}-\mathbf{GA})\mathbf{Z}\|&\qquad \text{すべての}\ \mathbf{Y}\in\mathscr{M}(\mathbf{A})\ \text{と任意の}\ \mathbf{Z}\ \text{にたいして}\\ \Leftrightarrow(\mathbf{GY},(\mathbf{I}-\mathbf{GA})\mathbf{Z})=0&\qquad \text{すべての}\ \mathbf{Y}\in\mathscr{M}(\mathbf{A})\ \text{と}\ \mathbf{Z}\ \text{にたいして}\\ \Leftrightarrow(\mathbf{GA})^*\mathbf{N}(\mathbf{I}-\mathbf{GA})=0\Leftrightarrow(\mathbf{GA})^*\mathbf{N}&=\mathbf{NGA}\end{aligned}$$

となり，これに $\mathbf{G}$ が一般逆行列である，すなわち，$\mathbf{AGA}=\mathbf{A}$ という条件を加えると，(1c. 5. 1) が証明される．

1つの選択は，次であることがわかる．

$$\mathbf{A}^-_{m(\mathbf{N})}=\mathbf{N}^{-1}\mathbf{A}^*(\mathbf{AN}^{-1}\mathbf{A}^*)^-. \tag{1c. 5. 2}$$

（ii）**最小2乗型一般逆行列**　$\mathbf{AX}=\mathbf{Y}$ は必ずしも解をもたない方程式，$\mathbf{G}$ は，$\mathbf{M}$ を p. d. 行列とするとき，$\mathbf{X}=\mathbf{GY}$ が $(\mathbf{Y}-\mathbf{AX})^*\mathbf{M}(\mathbf{Y}-\mathbf{AX})$ を最小にする行列とする．そのための必要十分条件は，

$$\mathbf{AGA}=\mathbf{A},\qquad (\mathbf{AG})^*\mathbf{M}=\mathbf{MAG} \tag{1c. 5. 3}$$

である．そのような $\mathbf{G}$ を $\mathbf{A}^-_{l(\mathbf{M})}$ であらわし，$\mathbf{M}=\mathbf{I}$ のとき，$\mathbf{A}^-_l$ と書く．

仮説によって，すべての $\mathbf{X},\mathbf{Y}$ にたいして，

$$(\mathbf{AGY}-\mathbf{Y})^*\mathbf{M}(\mathbf{AGY}-\mathbf{Y})\leqslant(\mathbf{AX}-\mathbf{Y})^*\mathbf{M}(\mathbf{AX}-\mathbf{Y})$$

つまり，すべての $\mathbf{X},\mathbf{Y}$ にたいして，

$$\|\mathbf{AGY}-\mathbf{Y}\|\leqslant\|\mathbf{AX}-\mathbf{Y}\|=\|\mathbf{AX}-\mathbf{AGY}+\mathbf{AGY}-\mathbf{Y}\|$$
$$\Leftrightarrow \mathbf{A}^*\mathbf{MAG}=\mathbf{A}^*\mathbf{M}\Leftrightarrow \mathbf{AGA}=\mathbf{A},\ (\mathbf{AG})^*\mathbf{M}=\mathbf{MAG}.$$

$\mathbf{G}$ として

$$\mathbf{A}^-_{l(\mathbf{M})}=(\mathbf{A}^*\mathbf{MA})^-\mathbf{A}^*\mathbf{M} \tag{1c. 5. 4}$$

を選び，最小解を与えるベクトル $\mathbf{X}$ を用いると，方程式

$$\mathbf{A}^*\mathbf{MAX}=\mathbf{A}^*\mathbf{MY}$$

は満足され，これにたいする一般解は $\mathbf{GY}+(\mathbf{I}-\mathbf{GA})\mathbf{Z}$ となり，$\mathbf{G}$ は(1c. 5. 3) を満足することがわかる．

（iii）**ノルム最小の最小2乗型一般逆行列**　$\mathbf{AX}=\mathbf{Y}$ は必ずしも解をもたない方程式，$\mathbf{G}$ は，$\mathbf{GY}$ が $(\mathbf{Y}-\mathbf{AX})^*\mathbf{M}(\mathbf{Y}-\mathbf{AX})$ を最小にする $\mathbf{X}$ の集合のなかで最小のノルム（(i) で定義したような）をもつ行列とする．このための必要十分条件は，

$$\mathbf{AGA}=\mathbf{A},\quad (\mathbf{AG})^*\mathbf{M}=\mathbf{MAG},\quad \mathbf{GAG}=\mathbf{G},\quad (\mathbf{GA})^*\mathbf{N}=\mathbf{NGA} \tag{1c. 5. 5}$$

である．そのような $\mathbf{G}$ を $\mathbf{A}^+_{\mathbf{MN}}$ であらわし，$\mathbf{M}=\mathbf{I}$，$\mathbf{N}=\mathbf{I}$ のとき $\mathbf{A}^+$ と書く．

$\mathbf{GY}$ が $(\mathbf{Y}-\mathbf{AX})^*\mathbf{M}(\mathbf{Y}-\mathbf{AX})$ を最小にするのだから，(1c. 5. 5) の最初の2つの条件は満足される．この性質をもつ $\mathbf{X}$ としての一般解は，(ii) に示したように，$\mathbf{GY}+(\mathbf{I}-\mathbf{GA})\mathbf{Z}$，$\mathbf{Z}$ は任意，である．このとき，

$$\|\mathbf{GY}\|\leqslant\|\mathbf{GY}+(\mathbf{I}-\mathbf{GA})\mathbf{Z}\|\qquad \text{すべての } \mathbf{Y},\mathbf{Z} \text{ について}$$
$$\Leftrightarrow (\mathbf{GY},(\mathbf{I}-\mathbf{GA})\mathbf{Z})=0,\ \text{すべての } \mathbf{Y},\mathbf{Z} \text{ について} \Leftrightarrow \mathbf{G}^*\mathbf{N}(\mathbf{I}-\mathbf{GA})=\mathbf{0}$$

を得，これは (1c. 5. 5) のあとの2つの条件と同等になる．

（iv）(iii) で定義した $\mathbf{A}^+_{\mathbf{MN}}$ は，$\mathbf{M}$ と $\mathbf{N}$ が p. d. 行列なら，一意である．

線形方程式のノルム最小解は一意であるから，この結果は正しい．

$$\mathbf{A}^+_{\mathbf{MN}}=\mathbf{N}^{-1}\mathbf{A}^*\mathbf{MA}(\mathbf{A}^*\mathbf{MAN}^{-1}\mathbf{A}^*\mathbf{MA})^-\mathbf{A}^*\mathbf{M} \tag{1c. 5. 6}$$

が条件 (1c. 5. 5) を満足することは容易に証明できる．$\mathbf{M}=\mathbf{I}$，$\mathbf{N}=\mathbf{I}$ という特別の場合の $\mathbf{A}^+$ は，すでに [(viii)，1b. 5] で，ムーア＝ペンローズ逆行列として紹介した．

ノルム最小型逆行列と最小2乗型逆行列の双対関係を示す重要な定理を (v) に与える．

（v）$\mathbf{M}$ を p. d. 行列，$\mathbf{A}$ を任意の $m\times n$ 行列とする．このとき，

$$(\mathbf{A}^*)^-_{m(\mathbf{M})}=[\mathbf{A}^-_{l(\mathbf{M}^{-1})}]^* \tag{1c. 5. 7}$$

(1c. 5. 7) の左辺を $\mathbf{G}$ とすると，定義により，

$$\mathbf{A}^*\mathbf{GA}^*=\mathbf{A}^*,\qquad (\mathbf{GA}^*)^*\mathbf{M}=\mathbf{M}(\mathbf{GA}^*),$$

これを書きかえると，$\mathbf{AG}^*\mathbf{A}=\mathbf{A}$，$(\mathbf{AG}^*)^*\mathbf{M}^{-1}=\mathbf{M}^{-1}(\mathbf{AG}^*)$ となり，つまり，$\mathbf{G}^*=\mathbf{A}^-_{l(\mathbf{M}^{-1})}$ を得る．したがって，$\mathbf{G}=[\mathbf{A}^-_{l(\mathbf{M}^{-1})}]^*$.

（vi）$\mathbf{M},\mathbf{N}$ を p. d. 行列とするとき，

（a）$(\mathbf{A}^+_{\mathbf{NM}})^+_{\mathbf{NM}}=\mathbf{A}$, (1c. 5. 8)

（b）$(\mathbf{A}^+_{\mathbf{MN}})^*=(\mathbf{A}^*)^+_{\mathbf{N}^{-1}\mathbf{M}^{-1}}$. (1c. 5. 9)

この結果は，$\mathbf{A}^+_{\mathbf{MN}}$ の定義から出る．

一般逆行列に関するこれ以上の結果については，Rao and Mitra (1971) を見よ．

1c. 6　制限づき固有値問題

ある種の問題では，ベクトル $\mathbf{X}$ に制約条件

$$\mathbf{X}'\mathbf{BX}=1,\qquad \mathbf{C}'\mathbf{X}=\mathbf{0} \tag{1c. 6. 1}$$

ただし，$\mathbf{B}$ は正定符号，$\mathbf{C}$ は $m\times k$ 行列，があるときの，2次形式 $\mathbf{X}'\mathbf{AX}$ の停留点およびその値を必要とする．$\mathbf{P}=\mathbf{C}(\mathbf{C}'\mathbf{B}^{-1}\mathbf{C})^-\mathbf{C}'\mathbf{B}^{-1}$ を $\mathscr{M}(\mathbf{C})$ の上への射影作用素とする．次の命題が得られる．

（i）$\mathbf{X}$ に制約 (1c. 6. 1) が伴なうときの $\mathbf{X}'\mathbf{AX}$ の停留点とその値は，$(\mathbf{I}-\mathbf{P})\mathbf{A}$ の $\mathbf{B}$ に関する固有ベクトルと固有値である．

ラグランジュ乗数 λ,μ を導入して，式

$$\mathbf{X}'\mathbf{AX}-\lambda(\mathbf{X}'\mathbf{BX}-1)-2\mathbf{X}'\mathbf{C}\mu$$

を考え，その導関数を零に等しいとおくと，

$$\begin{aligned}\mathbf{AX}-\lambda\mathbf{BX}-\mathbf{C}\mu&=\mathbf{0}\\ \mathbf{C}'\mathbf{X}&=\mathbf{0}\\ \mathbf{X}'\mathbf{BX}&=1.\end{aligned} \tag{1c. 6. 2}$$

を得る．(1c. 6. 2) の第1式に $\mathbf{I}-\mathbf{P}$ をかけると，

$$(\mathbf{I}-\mathbf{P})\mathbf{AX}-\lambda\mathbf{BX}=\mathbf{0}$$

を得，これが求める結果を証明する．これ以上の結果については，Rao (1964) を見よ．

注意　$\mathbf{A}=\mathbf{aa}'$，$\mathbf{a}$はm次元ベクトル，という特別の場合には，$\mathbf{X}'\mathbf{AX}$は，

$$\mathbf{X} \propto \mathbf{B}^{-1}(\mathbf{I}-\mathbf{P})\mathbf{a} \tag{1c.6.3}$$

のとき最大値をとる．これは重要な結果である．

(1c.6.1) において，条件 $\mathbf{C}'\mathbf{X}=\mathbf{0}$ を $\mathbf{C}'\mathbf{X}\geqslant\mathbf{0}$（つまり，$\mathbf{C}'\mathbf{X}$ の各要素が非負）におきかえる場合の問題は Rao (1964) で考えられている．

1d　ベクトル空間の凸集合

1d.1　定　義

ベクトル空間の部分集合 C が凸であるとは，任意の2点 $\mathbf{x}_1, \mathbf{x}_2 \in C$ にたいして，$\mathbf{x}_1$ と $\mathbf{x}_2$ を結ぶ線分が C に含まれる．つまり，$\lambda\mathbf{x}_1+(1-\lambda)\mathbf{x}_2 \in C$，$0\leqslant\lambda\leqslant 1$ の場合をいう．凸集合の例は，直線上の点，球や半空間内の点，などである．

任意の集合Aにたいして，Aに属する有限個の点の組から，凸結合

$$\lambda_1\mathbf{x}_1+\cdots+\lambda_k\mathbf{x}_k$$
$$\lambda_i\geqslant 0,\qquad \sum\lambda_i=1,\qquad \mathbf{x}_i\in A$$

によって生成される点の集合 A^* は，明らかに，A を含む凸集合になる．実際，A^* は A を含む最小の凸集合で，Aの凸包とよばれる．2点の凸包はその2点を結ぶ線分であり，球の凸包はその球と球内のすべての点になる．

内点，境界点，外点や，開集合と閉集合，および有界集合は，集合論におけると同様に定義する．すると，$\mathbf{x}$ が集合Aの内点であるとは，$\mathbf{x}$を中心とする球でAの部分集合であるようなものが存在することをいう．Aのすべての内点の集合は開集合であるという．その補集合が開集合であるような集合は閉集合である．A の境界点（A に属するとはかぎらない）とは，それを中心とするどの球もAの点とAの補集合 A^c の点を含むものをいう．集合の閉包は，その集合に含まれていなかったすべての境界点をその集合に付加することによって得られ，$\bar{A}$ であらわす．集合 $\bar{A}$ もまた A を含む最小の閉集合である．A の外点は $\bar{A}$ の補集合である $\bar{A}^c$ の点である．

凸集合の次の諸性質は容易に証明できる．(1)　任意の数の凸集合の交わりは凸集合である．(2)　凸集合の閉包は凸集合である．(3)　凸集合 C とその閉包 $\bar{C}$ は同じ内点，境界点，外点をもつ．(4)　$\mathbf{x}$をCの内点，$\mathbf{y}$ を境界点とすると，点 $(1-\lambda)\mathbf{x}+\lambda\mathbf{y}$ は $0\leqslant\lambda<1$ のときCの内点，$\lambda>1$ のときCの外点となる．

1d.2　凸集合の分離定理

(i)　Cを凸集合，$\mathbf{y}$を外点とする．このとき，

$$\inf_{\mathbf{x}\in C}(\mathbf{a},\mathbf{x})>(\mathbf{a},\mathbf{y}) \tag{1d.2.1}$$

となるような，$\|\mathbf{a}\|=1$ のベクトル $\mathbf{a}$ が存在する．

$\mathbf{b}$を

$$\inf_{\mathbf{x}\in C}\|\mathbf{y}-\mathbf{x}\|=\|\mathbf{b}-\mathbf{y}\|$$

であるような境界点とする．$\mathbf{x}\in C$ なら，$0\leqslant\lambda\leqslant 1$ にたいして，点 $(1-\lambda)\mathbf{x}+\lambda\mathbf{b}\in\bar{C}$ であるから，

$$\|(1-\lambda)\mathbf{x}+\lambda\mathbf{b}-\mathbf{y}\|\geqslant\|\mathbf{b}-\mathbf{y}\| \tag{1d.2.2}$$

となる．(1d.2.2) を展開すると，すべての $\lambda\leqslant 1$ にたいして，

$$(1-\lambda)^2\|\mathbf{x}-\mathbf{b}\|^2+2(1-\lambda)(\mathbf{x}-\mathbf{b},\mathbf{b}-\mathbf{y})\geqslant 0$$

を得る．したがって，$(\mathbf{x}-\mathbf{b},\mathbf{b}-\mathbf{y})\geqslant 0$ となり，

$$(\mathbf{x},\mathbf{b}-\mathbf{y})\geqslant(\mathbf{b},\mathbf{b}-\mathbf{y})>(\mathbf{y},\mathbf{b}-\mathbf{y})$$

を得る．ここで，$\mathbf{a}=\mathbf{b}-\mathbf{y}$ にとると，(1d.2.1) の結果が証明される．また，確かに $\mathbf{a}$ を規準化することもできる．

(ii)　Cを凸集合，$\mathbf{y}$をCの境界上の点とする．このとき，$\mathbf{y}$を通る支持平面が存在する．すなわち，

$$\inf_{\mathbf{x}\in C}(\mathbf{a},\mathbf{x})=(\mathbf{a},\mathbf{y})\quad\text{つまり，}$$
$$(\mathbf{a},\mathbf{x})\geqslant(\mathbf{a},\mathbf{y})\quad\text{すべての }\mathbf{x}\in C\text{ にたいして} \tag{1d.2.3}$$

とするような非零ベクトル $\mathbf{a}$ が存在する．

$\mathbf{y}_n$ を $\lim\mathbf{y}_n=\mathbf{y}$ となるような外点の系列とする．(1d.2.1) によって $\mathbf{y}_n$ から定められる $\mathbf{a}_n$ を考える．$\|\mathbf{a}_n\|$ は有界であるから，極値 $\mathbf{a}$ に収束するような部分系列 $\mathbf{a}_\nu$ が存在する．したがって，

$$(\mathbf{x},\mathbf{a})=\lim_{\nu\to\infty}(\mathbf{x},\mathbf{a}_\nu)\geqslant\lim_{\nu\to\infty}(\mathbf{y}_\nu,\mathbf{a}_\nu)=(\mathbf{y},\mathbf{a}).$$

(iii)　CとDを内点をまったく共有しない2つの凸集合とする．このとき，ベクトル $\mathbf{a}$ とスカラー α が存在して，すべての $\mathbf{x}\in C$ にたいし $(\mathbf{a},\mathbf{x})\geqslant\alpha$ で，すべての $\mathbf{y}\in D$ にたいして $(\mathbf{a},\mathbf{y})\leqslant\alpha$ とできる．すなわち，2つの集合を分離する超平面が存在する．

$\mathbf{x}\in C$，$\mathbf{y}\in D$ にたいして点 $\mathbf{x}-\mathbf{y}$ の凸集合を考えると，これは0を内点として含まない．(1d.2.1) と (1d.2.3) によって，

$$(\mathbf{a},\mathbf{x}-\mathbf{y})\geqslant(\mathbf{a},\mathbf{0})=0\qquad\text{つまり}\qquad(\mathbf{a},\mathbf{x})\geqslant(\mathbf{a},\mathbf{y})$$

とするような $\mathbf{a}$ が存在する．したがって，

$$(\mathbf{a},\mathbf{x})\geqslant\alpha,\ \mathbf{x}\in C\qquad\text{かつ}\qquad(\mathbf{a},\mathbf{y})\leqslant\alpha,\ \mathbf{y}\in D \tag{1d.2.4}$$

とするような α が存在する．

C と D を共有点をまったくもたない閉凸集合とすると，$\mathbf{0}$ は点 $\mathbf{x}-\mathbf{y}$ の閉集合の外点となり，この場合は，(1d. 2. 1) の適用によって，

$$(\mathbf{a},\mathbf{x})>\alpha,\ \mathbf{x}\in C \quad かつ \quad (\mathbf{a},\mathbf{y})<\alpha,\ \mathbf{y}\in D \tag{1d. 2. 5}$$

という厳密な分離が得られる．

ベクトル $\mathbf{x}\in E_n$ は，そのすべての座標が非負で，少なくとも1つの座標は正であるとき $>\mathbf{0}$ という．負ベクトルはすべての座標が負のものをいう．

(iv)　C は E_n の凸集合でその点はどれも負でないとき，すべての $\mathbf{x}\in C$ にたいして，$(\mathbf{a},\mathbf{x})>0$ となるようなベクトル $\mathbf{a}>\mathbf{0}$ が存在する．

凸集合 C とすべての負のベクトルの集合 Y とを考える．(iii) の結果から，

$$(\mathbf{a},\mathbf{x})\geqslant\alpha,\ \mathbf{x}\in C \quad かつ \quad (\mathbf{a},\mathbf{y})\leqslant\alpha,\ \mathbf{y}\in Y \tag{1d. 2. 6}$$

となるような $\mathbf{a}$ と α が存在する．$\mathbf{y}\to\mathbf{0}$ のとき $(\mathbf{a},\mathbf{y})\to 0$ となるから，$\alpha\geqslant 0$ である．さらに，$\mathbf{a}>\mathbf{0}$ である．そうでなければ，$(\mathbf{a},\mathbf{y})\geqslant\alpha$ となるような負のベクトル $\mathbf{y}$ をみつけることができ，これは (1d. 2. 6) と矛盾する．したがって，$(\mathbf{a},\mathbf{x})>0$ とするような $\mathbf{a}>\mathbf{0}$ が存在する．

凸集合の簡単な取り扱いは Berge (1963) の本にみられる．読者は次の結果を容易に確かめることができるであろう．

(v)　**ファルカスの補助定理**　β と $\alpha_1,\cdots,\alpha_k$ は R^n のベクトルとし，$\mathbf{x}$ はすべての i にたいして $\alpha_i'\mathbf{x}\geqslant 0$ となるとする．このとき，$\alpha'\mathbf{x}\geqslant 0$ となるようなすべての $\mathbf{x}$ にたいして $\beta'\mathbf{x}\geqslant 0$ となる必要十分条件は，$a_i\geqslant 0$ にたいして $\beta=\sum a_i\alpha_i$ と書けることである．

証明は省略する．

1. E_n の部分集合 C の凸包を C^* とすると，C^* の各点は C の高々 $n+1$ 個の点の凸結合としてあらわされる．
2. E_n の有界集合を C とし，
$$\mathbf{x}=\sum_{i=1}^{\infty}\alpha_i\mathbf{x}_i,\qquad \mathbf{x}_i\in C,\quad \alpha_i>0,\quad \sum\alpha_i=1$$
とすると，$\mathbf{x}\in C^*$ となる．
3. 点 $\mathbf{x}\in C$ が C の端点であるとは，ある λ $(0<\lambda<1)$ にたいして，異なる2つの点 $\mathbf{x}_1,\mathbf{x}_2\in C$ によって $\mathbf{x}=\lambda\mathbf{x}_1+(1-\lambda)\mathbf{x}_2$ と書けないときをいう．C が有界閉凸集合なら，C は，その端点によって張られる．すなわち，どの $\mathbf{x}\in C$ も，C の端点 $\mathbf{x}_1,\cdots,\mathbf{x}_k$ を用いて，
$$\mathbf{x}=\sum_{1}^{k}\alpha_i\mathbf{x}_i,\qquad \alpha_i\geqslant 0,\qquad \sum\alpha_i=1$$
という形であらわされる．
4. 有界閉凸集合はどの支持超平面においても端点をもつ．
5. E_n の閉凸集合は，有限個の閉半空間の交わりであるときのみ，有限個の端点をもつ．

1e 不等式

ベクトル，行列，行列式に関するいくつかの標準的な不等式は数理統計学への応用が多いので，本節に集めておく．

1e. 1 コーシー＝シュヴァルツ (C-S) 不等式

(i)　実要素をもつ任意の2つの列ベクトル $\mathbf{X},\mathbf{Y}$ について，

$$(\mathbf{X}'\mathbf{Y})^2\leqslant(\mathbf{X}'\mathbf{X})(\mathbf{Y}'\mathbf{Y}),$$

これを $\mathbf{X}$ と $\mathbf{Y}$ の要素で書くと

$$(\sum x_iy_i)^2\leqslant(\sum x_i^2)(\sum y_i^2), \tag{1e. 1. 1}$$

ここで，等号は，ある実スカラー λ,μ にたいして $\lambda\mathbf{X}+\mu\mathbf{Y}=\mathbf{0}$ が成り立つときにのみ得られる．

λ と μ の2次形式

$$(\lambda\mathbf{X}+\mu\mathbf{Y})'(\lambda\mathbf{X}+\mu\mathbf{Y})=\lambda^2\mathbf{X}'\mathbf{X}+2\lambda\mu\mathbf{X}'\mathbf{Y}+\mu^2\mathbf{Y}'\mathbf{Y}$$

は非負である．したがって，[(v), **1c. 1**] または初等代数学を使うと，この2次形式の判別式は正か零でなければならない．すなわち

$$\begin{vmatrix}\mathbf{X}'\mathbf{X} & \mathbf{Y}'\mathbf{X}\\ \mathbf{X}'\mathbf{Y} & \mathbf{Y}'\mathbf{Y}\end{vmatrix}\geqslant 0,\qquad よって\ (\mathbf{X}'\mathbf{Y})^2\leqslant(\mathbf{X}'\mathbf{X})(\mathbf{Y}'\mathbf{Y}). \tag{1e. 1. 2}$$

等号は判別式が零のときに得られ，そのときは，第1行の λ 倍と第2行の μ 倍の和が零になるような λ と μ が存在することになる．したがって $(\lambda\mathbf{X}+\mu\mathbf{Y})'(\lambda\mathbf{X}+\mu\mathbf{Y})=0$，よって $\lambda\mathbf{X}+\mu\mathbf{Y}=\mathbf{0}$ を得る．

C-S 不等式の他の2つの形式は興味があるので，次に示す．

(ii)　$\mathbf{X},\mathbf{Y}$ を2つの列ベクトル，$\mathbf{A}$ をグラム行列すなわち $\mathbf{A}=\mathbf{B}'\mathbf{B}$ とする．このとき

(a)　$(\mathbf{X}'\mathbf{A}\mathbf{Y})^2\leqslant(\mathbf{X}'\mathbf{A}\mathbf{X})(\mathbf{Y}'\mathbf{A}\mathbf{Y})$，　等号は $\mathbf{X}\propto\mathbf{Y}$ のとき成り立つ．　(1e. 1. 3)

(b)　$(\mathbf{X}'\mathbf{Y})^2\leqslant(\mathbf{X}'\mathbf{A}\mathbf{X})(\mathbf{Y}'\mathbf{A}^{-1}\mathbf{Y})$，　(1e. 1. 4)

$\mathbf{A}^{-1}$ が存在するとき (b) が得られ，その等号は $\mathbf{X}\propto\mathbf{A}^{-1}\mathbf{Y}$ のとき成り立つ．

$\mathbf{U}=\mathbf{B}\mathbf{X}$，$\mathbf{V}=\mathbf{B}\mathbf{Y}$ にとり，$\mathbf{U}$ と $\mathbf{V}$ に C-S 不等式を適用すれば (a) の結果が得られる．同様に，(b) は

$$\mathbf{U}=(\mathbf{B}^{-1})'\mathbf{Y},\qquad \mathbf{V}=\mathbf{B}\mathbf{X}$$

にとればよい．

（ⅲ）C-S 不等式を積分に変形すると

$$\left(\int_A fgdv\right)^2 \leqslant \int_A f^2dv\int_A g^2dv \tag{1e. 1. 5}$$

となる．ここで，f と g はある集合 A で定義された実関数で，f^2, g^2 は測度 v に関して積分可能とする．

この結果を証明するためには，λ と μ の非負2次形式，

$$\int_A(\lambda f+\mu g)^2dv = \lambda^2\int_A f^2dv + 2\lambda\mu\int_A fgdv + \mu^2\int_A g^2dv \tag{1e. 1. 6}$$

を考えればよい．特に，v が確率測度のときには，x, y を確率変数，cov，V をそれぞれ共分散と分散をあらわすとすると，

$$[\mathrm{cov}(xy)]^2 \leqslant V(x)V(y) \tag{1e. 1. 7}$$

を得る．

ラグランジュの等式　不等式 (1e. 1. 1) は，x_i, y_i をそれぞれベクトル $\mathbf{X}, \mathbf{Y}$ の成分とすると，等式

$$(\sum x_i^2)(\sum y_i^2) - (\sum x_iy_i)^2 = \sum_{i<j}(x_iy_j - x_jy_i)^2 \tag{1e. 1. 8}$$

からただちに導ける．

1e. 2　ヘルダーの不等式

$x_i, y_i \geqslant 0$，$i = 1, \cdots, n$ で $(1/p) + (1/q) = 1$，$p > 1$ のとき，

$$\sum_1^n x_iy_i \leqslant \left(\sum_1^n x_i^p\right)^{1/p}\left(\sum_1^n y_i^q\right)^{1/q} \tag{1e. 2. 1}$$

で，等号は $y_i \propto x_i^{p-1}$ のときのみ成立する．これを積分であらわすと，$f, g \geqslant 0$ にたいして，

$$\int_R fgdv \leqslant \left(\int_R f^pdv\right)^{1/p}\left(\int_R g^qdv\right)^{1/q}.$$

まず，微分または他の方法で，

$$\min_{x>0}\left\{t(x) = \frac{x^p}{p} + \frac{x^{-q}}{q}\right\} = 1, \qquad x = 1 \text{ のとき}$$

となることを証明する．$t(x) \geqslant 1$ に $x = u^{1/q}v^{-1/p}$ $(u, v \geqslant 0)$ を代入すると

$$uv \leqslant \frac{u^p}{p} + \frac{v^q}{q} \tag{1e. 2. 2}$$

を得，等号は $v = u^{p-1}$ のとき成り立つことがわかる．

$u_k = x_k/(\sum x_k^p)^{1/p}$，$v_k = y_k/(\sum y_k^q)^{1/q}$ とせよ．これを (1e. 2. 2) に代入し，k について加えると

$$\sum u_kv_k \leqslant \frac{\sum u_k^p}{p} + \frac{\sum v_k^q}{q} = \frac{1}{p} + \frac{1}{q} = 1$$

を得，求める不等式が得られる．その積分形も同様に証明できる．より一般的には，次の不等式を得る．

$$\sum x_iy_iz_i\cdots \leqslant (\sum x_i^p)^{1/p}(\sum y_i^q)^{1/q}(\sum z_i^r)^{1/r}\cdots \quad \text{ただし} \quad \frac{1}{p}+\frac{1}{q}+\frac{1}{r}\cdots \leqslant 1. \tag{1e. 2. 3}$$

ミンコフスキの不等式　ヘルダーの不等式から次を導くことは，容易である．$x_i, y_i \geqslant 0$ で $k \geqslant 1$ なら

$$[\sum(x_i+y_i)^k]^{1/k} \leqslant (\sum x_i^k)^{1/k} + (\sum y_i^k)^{1/k}. \tag{1e. 2. 4}$$

1e. 3　アダマールの不等式

正則実 $n \times n$ 行列 $\mathbf{B} = (b_{ij})$ にたいして

$$|\mathbf{B}|^2 \leqslant \prod_{i=1}^n\left(\sum_{k=1}^n b_{ik}^2\right) \tag{1e. 3. 1}$$

が成り立つ．特別の場合として，

（ⅰ）$\sum b_{ik}^2 = 1$，$i = 1, \cdots, n$ なら，$|\mathbf{B}| \leqslant 1$　(1e. 3. 2)

（ⅱ）$|b_{ij}| \leqslant M$ なら，$|\mathbf{B}| \leqslant M^nn^{n/2}$.　(1e. 3. 3)

(1e. 3. 1) を証明するために，正定符号行列 $\mathbf{A} = \mathbf{BB}'$ を考え，不等式 $|\mathbf{A}| \leqslant a_{11}\cdots a_{nn}$（対角要素の積）を適用する．展開

$$|\mathbf{A}| = a_{11}\begin{vmatrix} a_{22} & \cdots & a_{2n} \\ & \cdots\cdots\cdots & \\ a_{n2} & \cdots & a_{nn}\end{vmatrix} + \begin{vmatrix} 0 & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & \cdots\cdots\cdots\cdots & & \\ a_{n1} & a_{n2} & \cdots & a_{nn}\end{vmatrix}$$

を考える．$\mathbf{A}$ は p. d. であるから，行列 (a_{ij})，$i, j = 2, \cdots, n$ もまた p. d. であり，上の第2項は負である．とすると，

$$|\mathbf{A}| \leqslant a_{11}\begin{vmatrix} a_{22} & \cdots & a_{2n} \\ & \cdots\cdots\cdots & \\ a_{n2} & \cdots & a_{nn}\end{vmatrix}$$

であり，$|\mathbf{A}| \leqslant a_{11}\cdots a_{nn}$ という結果は帰納法によって求められる．(1e. 3. 1) の結果は，$a_{ii} = \sum b_{ik}^2$ であることに注意するとすぐに出る．

1e. 4　積率に関する不等式

x を，$E(x) = \mu$ と積率

$$E(x-\mu)^r = \mu_r, \qquad r = 1, \cdots, 2N$$

が存在するような確率変数とする．y_i の2次形式 $\sum\sum\mu_{i+j-2}y_iy_j$ は

$$E[y_1+y_2(x-\mu)+\cdots+y_{N+1}(x-\mu)^N]^2\geqslant 0 \qquad \text{(1e. 4. 1)}$$

となる．したがって，(μ_{i+j-2}) は非負定符号である．これは，しかしながら，μ_r がある確率変数の積率系列であるための十分条件ではない．これは，その積率が離散または連続分布のものであっても，観測値の所与の組から計算したものであっても，必要条件である．特に，$N=2$ のときは，次を得る，

$$\begin{vmatrix}1 & & \mu_2\\ 0 & \mu_2 & \mu_3\\ \mu_2 & \mu_3 & \mu_4\end{vmatrix}\geqslant 0 \quad \text{よって，} \quad \mu_2^3\left(\frac{\mu_4}{\mu_2^2}-\frac{\mu_3^2}{\mu_2^3}-1\right)\geqslant 0$$

$$\text{すなわち，}\ \beta_2-\beta_1-1\geqslant 0. \qquad \text{(1e. 4. 2)}$$

ここで，$\beta_1=\mu_3^2/\mu_2^3$ と $\beta_2=\mu_4/\mu_2^2$ は，分布の非対称性ととがりをあらわす尺度である．

1e. 5　凸関数とイェンセンの不等式

関数 $f(x)$ が凸であるとは，$\alpha,\beta>0$，$\alpha+\beta=1$ として，

$$f(\alpha x+\beta y)\leqslant \alpha f(x)+\beta f(y) \qquad \text{すべての } x,y \text{ にたいして} \qquad \text{(1e. 5. 1)}$$

が成り立つときを言う．そのような関数について，まず，任意の点 x_0 において，右および左導関数 $f'_+(x_0)$，$f'_-(x_0)$ が存在することを証明しよう．$x_0<x_1<x_2$ とする．$\alpha=(x_2-x_1)/(x_2-x_0)$，$\beta=(x_1-x_0)/(x_2-x_0)$，$x=x_0$，$y=x_2$ を $\alpha x+\beta y=x_1$ となるように選び，(1e. 5. 1) を用いると，x_2-x_0 をかけて，

$$(x_2-x_0)f(x_1)\leqslant (x_2-x_1)f(x_0)+(x_1-x_0)f(x_2) \qquad \text{(1e. 5. 2)}$$

が成立することがわかる．(1e. 5. 2) の両辺に $x_0f(x_0)$ を加え，移項すると，

$$\frac{f(x_1)-f(x_0)}{x_1-x_0}\leqslant\frac{f(x_2)-f(x_0)}{x_2-x_0} \qquad \text{(1e. 5. 3)}$$

を得，これから $[f(x)-f(x_0)]/(x-x_0)$ は $x\to x_0$ とともに減少することがわかる．(1e. 5. 2) の両辺に $x_1f(x_1)$ を加え，移項すると，

$$\frac{f(x_0)-f(x_1)}{x_0-x_1}\leqslant\frac{f(x_2)-f(x_1)}{x_2-x_1} \qquad \text{(1e. 5. 4)}$$

を得る．式 (1e. 5. 4) を，$x_{-1}<x_0<x_1$ の項であらわすと，

$$\frac{f(x_{-1})-f(x_0)}{x_{-1}-x_0}\leqslant\frac{f(x_1)-f(x_0)}{x_1-x_0}$$

となり，これから，$[f(x)-f(x_0)]/(x-x_0)$ は下に有界であり，$x\to x_0$ とともに減少するから，右導関数 $f'_+(x_0)$ が存在することが示される．同様に，$f'_-(x_0)$ が存在し，明らかに $f'_-(x_0)\leqslant f'_+(x_0)$ である．L を $f'_-(x_0)\leqslant L\leqslant f'_+(x_0)$ とする．このとき，すべての x にたいして，

$$f(x)\geqslant f(x_0)+L(x-x_0) \qquad \text{(1e. 5. 5)}$$

が成り立つ．それは，$x_1>x_0$ なら

$$\frac{f(x_1)-f(x_0)}{x_1-x_0}\geqslant f'_+(x_0)\geqslant L$$

だからである．また，$x_1<x_0$ のときには逆向きの関係が成り立つ．不等式 (1e. 5. 5) は重要な応用をもつ．すなわち，それからイェンセンの不等式が導かれる．

イェンセンの不等式　x を $E(x)=\mu$ なる確率変数，$f(x)$ を凸関数とすると，

$$E[f(x)]\geqslant f[E(x)] \qquad \text{(1e. 5. 6)}$$

で，等号は，〔$f(x)$ が線形の場合かまたは〕x が μ で退化した分布のときのみ成立する．

式 (1e. 5. 5) で x_0 のかわりに μ とおいた式，

$$f(x)\geqslant f(\mu)+L(x-\mu) \qquad \text{(1e. 5. 7)}$$

を考え，この両辺の期待値をとる．(1e. 5. 7) の右辺の第2項の期待値は零となるから，不等式 (1e. 5. 6) が得られる．

1e. 6　情報理論における不等式

（i）$\sum a_i$ と $\sum b_i$ は正の数の収束系列で $\sum a_i\geqslant\sum b_i$ とする．このとき，

$$\sum a_i\log\frac{b_i}{a_i}\leqslant 0 \qquad \text{(1e. 6. 1)}$$

で，等号は $a_i=b_i$ のときのみ成立する．さらに，すべての i について，$a_i\leqslant 1$，かつ $b_i\leqslant 1$ のときは，

$$2\sum a_i\log\frac{a_i}{b_i}\geqslant\sum a_i(a_i-b_i)^2. \qquad \text{(1e. 6. 2)}$$

これらの不等式を証明するために，$x>0$ のとき，$\log x$ の $x=1$ における展開は

$$\log x=(x-1)-(x-1)^2(2y^2)^{-1}, \qquad y\in(1,x) \qquad \text{(1e. 6. 3)}$$

となることに注目する．(1e. 6. 1) の各項に (1e. 6. 3) の展開を適用すると，

$$\sum a_i\log\frac{b_i}{a_i}=(\sum b_i-\sum a_i)-\sum a_i(b_i-a_i)^2(2y_i^2)^{-1}\leqslant 0 \qquad \text{(1e. 6. 4)}$$

となり，(1e. 6. 1) が証明される．(1e. 6. 4) では，$y_i^2\in(a_i^2,b_i^2)$ で，もし $a_i\leqslant 1$，$b_i\leqslant 1$ なら，y_i^2 の最大値は1以下となる．したがって，

$$-\sum\frac{(b_i-a_i)^2}{a_i}(2y_i^2)^{-1}\leqslant-\frac{1}{2}\sum a_i(b_i-a_i)^2 \qquad \text{(1e. 6. 5)}$$

となる．(1e. 6. 4) と (1e. 6. 5) を結合すると，(1e. 6. 2) を得る．

(1e. 6. 1) と (1e. 6. 2) の結果は，a_i と b_i が非負で，和は $a_i>0$ なる i についてのみとるとしても成立する．いいかえると，b_i のいくつかは零になってもよい（しかし，a_i は零ではいけない）のである．

（ii）f と g を非負で測度 μ に関して積分可能な関数，S を $f>0$ なる領域とする．

もし $\int_S (f-g)d\mu \geqslant 0$ なら，

$$\int_S f \log\frac{f}{g} d\mu \geqslant 0 \tag{1e.6.6}$$

で，等号は $f=g$(a. e. μ) のときのみ成立する．〔a. e. は付録 2B，127 ページ参照〕

この証明は，和を積分におきかえると，(1e. 6. 1) の証明と同じである．読者は，イェンセンの不等式 (1e. 5. 6) を適用すると，不等式 (1e. 6. 1) と (1e. 6. 6) を直接導くことができる．

1e. 7　スターリングの近似

これは階乗の便利な近似で，大きな数の階乗を実際に計算するときや，2項，ポアソン超幾何分布などの極限形を研究するときにも有用である．その近似は

$$n! \doteq \sqrt{2\pi n}\, n^n e^{-n}$$

と書かれ，これから，

$$\lim_{n\to\infty}\frac{1\cdot 2\cdot\cdots\cdot n}{\sqrt{2\pi n}\, n^n e^{-n}} = 1$$

が導ける．有限の n にたいしては，より正確な表現，

$$n! = \sqrt{2\pi n}\, n^n e^{-n} e^{\omega(n)/12}$$

がある．ただし，$\left(n+\frac{1}{2}\right)^{-1} < \omega(n) < n^{-1}$ である．近似 $\sqrt{2\pi n}\, n^n e^{-n}$ の絶対誤差は n が大きくなるにつれて大きくなるが，その相対誤差は急速に零に近づくことを注意しておこう．この公式は，大きな数の階乗の比を評価しなければならないときに有用である．証明は，数値計算の本を参照されたい．

1f　2次形式の極値

本節では，変数に1次および2次の制約があるときの2次形式の極値に関して，多くの有用な結果を示す．

1f. 1　一般的な結果

(i) $\mathbf{A}$ を $m\times m$ の p. d. 行列，$\mathbf{U}\in E_m$，つまり，$\mathbf{U}$ は次数 m のベクトルとする．このとき，

$$\sup_{\mathbf{X}\in E_m}\frac{(\mathbf{U}'\mathbf{X})^2}{\mathbf{X}\mathbf{A}'\mathbf{X}} = \mathbf{U}'\mathbf{A}^{-1}\mathbf{U} \tag{1f. 1. 1}$$

で，上限は $\mathbf{X}_* = \mathbf{A}^{-1}\mathbf{U}$ のとき得られる．

(1f. 1. 1)の結果は，C-S 不等式の1つの変形である (1e. 1. 4) を言いかえたにすぎない．

(i) の系として，任意の $m\times m$ 行列 $\mathbf{T}$ にたいして

$$\sup_{\mathbf{X}\in E_m}\frac{(\mathbf{U}'\mathbf{T}\mathbf{X})^2}{\mathbf{X}'\mathbf{A}\mathbf{X}} = \mathbf{U}'\mathbf{T}\mathbf{A}^{-1}\mathbf{T}'\mathbf{U}. \tag{1f. 1. 2}$$

(ii) $\mathbf{A}$ を $m\times m$ の p. d. 行列，$\mathbf{B}$ を $m\times k$ 行列，$\mathbf{U}$ を次数 k のベクトルとする．$\mathbf{B}'\mathbf{A}^{-1}\mathbf{B}$ の任意の一般逆行列を $\mathbf{S}^-$ であらわす．このとき $\mathbf{X}$ を列ベクトルとすると，

$$\inf_{\mathbf{B}'\mathbf{X}=\mathbf{U}} \mathbf{X}'\mathbf{A}\mathbf{X} = \mathbf{U}'\mathbf{S}^-\mathbf{U} \tag{1f. 1. 3}$$

で，下限は $\mathbf{X}_* = \mathbf{A}^{-1}\mathbf{B}\mathbf{S}^-\mathbf{U}$ のとき得られる．

この問題は，当然解が存在しなければならないような1次制約 $\mathbf{B}'\mathbf{X}=\mathbf{U}$ を $\mathbf{X}$ が受けるときの，$\mathbf{X}$ の2次形式の最小を求めるものであることに注意する．

$\mathbf{B}'\mathbf{X}=\mathbf{U}$ が解をもつなら，$\boldsymbol{\lambda}\in E_k$ が存在して，$\mathbf{B}'\mathbf{A}^{-1}\mathbf{B}\boldsymbol{\lambda}=\mathbf{U}$，すなわち，$\mathbf{S}\boldsymbol{\lambda}=\mathbf{U}$ とできる．その結果，

$$\mathbf{B}'\mathbf{X}_* = \mathbf{B}'\mathbf{A}^{-1}\mathbf{B}\mathbf{S}^-\mathbf{U} = \mathbf{S}\mathbf{S}^-\mathbf{S}\boldsymbol{\lambda} = \mathbf{S}\boldsymbol{\lambda} = \mathbf{U}$$

$$\mathbf{X}_*'\mathbf{A}\mathbf{X}_* = (\mathbf{U}'\mathbf{S}^{-\prime}\mathbf{B}'\mathbf{A}^{-1})\mathbf{A}(\mathbf{A}^{-1}\mathbf{B}\mathbf{S}^-\mathbf{U}) = \mathbf{U}'\mathbf{S}^-\mathbf{U} \tag{1f. 1. 4}$$

を得る．われわれは，任意の $\mathbf{X}$ にたいして，$\mathbf{X}'\mathbf{A}\mathbf{X} \geqslant \mathbf{U}'\mathbf{S}^-\mathbf{U}$ であることを示しさえすればよいが，それは，ベクトル $\mathbf{Y}_1 = \mathbf{A}^{1/2}\mathbf{X}$，$\mathbf{Y}_2 = \mathbf{A}^{1/2}\mathbf{X}_*$ に C-S 不等式を適用し，$\mathbf{Y}_1'\mathbf{Y}_2 = \mathbf{X}'\mathbf{B}\mathbf{S}^-\mathbf{U} = \mathbf{U}'\mathbf{S}^-\mathbf{U}$ となることを見ればすぐにわかる．(ii) の系として，所与の $\boldsymbol{\xi}\in E_m$ にたいし，

$$\inf_{\mathbf{B}'\mathbf{X}=\mathbf{U}} (\mathbf{X}-\boldsymbol{\xi})'\mathbf{A}(\mathbf{X}-\boldsymbol{\xi}) = (\mathbf{U}-\mathbf{B}'\boldsymbol{\xi})'\mathbf{S}^-(\mathbf{U}-\mathbf{B}'\boldsymbol{\xi}) \tag{1f. 1. 5}$$

で，その下限は $\mathbf{X}_* = \boldsymbol{\xi} + \mathbf{A}^{-1}\mathbf{B}\mathbf{S}^-\mathbf{U} - \mathbf{A}^{-1}\mathbf{B}\mathbf{S}^-\mathbf{B}'\boldsymbol{\xi}$ においてとられる．この結果は，$\mathbf{Y} = \mathbf{X}-\boldsymbol{\xi}$ とおき，(1f. 1. 3) を適用すると得られる．

(1f. 1. 3) の特殊の場合として，

$$\inf_{\mathbf{B}'\mathbf{X}=\mathbf{U}} \mathbf{X}'\mathbf{X} = \mathbf{U}'\mathbf{S}^-\mathbf{U} \tag{1f. 1. 6}$$

を得る．ここで，$\mathbf{S}^-$ は $\mathbf{B}'\mathbf{B}$ の g 逆行列で，下限は $\mathbf{X}_* = \mathbf{B}\mathbf{S}^-\mathbf{U}$ でとられる．

(1f. 1. 3)，(1f. 1. 5)，(1f. 1. 6) の結果で，$\mathbf{B}$ の階数が完全，すなわち，$\mathbf{B}$ の行の数 k に等しいときは，$\mathbf{S}^-$ は真の逆行列 $\mathbf{S}^{-1}$ になることに注意せよ．

(1f. 1. 6) の重要な一般化は，$\mathbf{X}_1$ を $\mathbf{X}$ の一部，たとえば，$\mathbf{X}$ の最初の q ($<k$) 個の成分のベクトルとするとき，$\mathbf{B}'\mathbf{X}=\mathbf{U}$ の条件の下で，$\mathbf{X}_1'\mathbf{X}_1$ の最小を求める問題である．このとき，条件 $\mathbf{B}'\mathbf{X}=\mathbf{U}$ は，次の分割形に書ける．

$$(\mathbf{B}_1' \,\vdots\, \mathbf{B}_2')\begin{pmatrix}\mathbf{X}_1\\ \cdots\\ \mathbf{X}_2\end{pmatrix} = \mathbf{U}. \tag{1f. 1. 7}$$

いま，任意の一般逆行列を選んで，次のようにおく．

$$\begin{pmatrix}\mathbf{B}_1'\mathbf{B}_1 & \mathbf{B}_2'\\ \mathbf{B}_2 & 0\end{pmatrix}^- = \begin{pmatrix}\mathbf{C}_1 & \mathbf{C}_2\\ \mathbf{C}_3 & \mathbf{C}_4\end{pmatrix} \tag{1f. 1. 8}$$

（iii）$\mathbf{B}_1, \mathbf{B}_2, \mathbf{C}_1$ を（1f. 1. 7）と（1f. 1. 8）で定義されるような行列とする．このとき

$$\inf_{\mathbf{B}'\mathbf{X}=\mathbf{U}} \mathbf{X}_1'\mathbf{X}_1 = (\mathbf{X}_{1*})'(\mathbf{X}_{1*}) = \mathbf{U}'\mathbf{C}_1\mathbf{U} \tag{1f. 1. 9}$$

である．ただし，$\mathbf{X}_{1*} = \mathbf{B}_1\mathbf{C}_1'\mathbf{U}$.

条件（1f. 1. 7）から，

$$\mathbf{U} = \mathbf{B}_1'\mathbf{B}_1\mathbf{E} + \mathbf{B}_2'\mathbf{F}, \qquad \mathbf{0} = \mathbf{B}_2\mathbf{E}$$

となるような列ベクトル $\mathbf{E}$ と $\mathbf{F}$ が存在することがわかる．このとき，

$$\mathbf{X}_{1*} = \mathbf{B}_1\mathbf{C}_1'\mathbf{U}, \qquad \mathbf{X}_{2*} = \mathbf{C}_2'\mathbf{U}$$

が，条件 $\mathbf{B}'\mathbf{X} = \mathbf{U}$ と $(\mathbf{X}_{1*})'(\mathbf{X}_{1*}) = \mathbf{U}'\mathbf{C}_1\mathbf{U}$ を満足することは容易に証明できる．さらに，$\mathbf{X}_1$ を条件（1f. 1. 7）を満足する他の任意のベクトルとしても $(\mathbf{X}_{1*})'\mathbf{X}_1 = \mathbf{U}'\mathbf{C}_1\mathbf{U}$ が成り立つ．（1f. 1. 9）の結果は，ベクトル $\mathbf{X}_1$ と $\mathbf{X}_{1*}$ に C-S 不等式を適用しても得られる．

（iv）$\mathbf{A}$ を次数 m の n. n. d. 行列，$\mathbf{B}$ を $m \times k$ 行列，$\mathbf{U}$ を $\mathbf{U} \in \mathscr{M}(\mathbf{B}')$ なる次数 k のベクトルとする．さらに，

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}' & \mathbf{0} \end{pmatrix}^- = \begin{pmatrix} \mathbf{C}_1 & \mathbf{C}_2 \\ \mathbf{C}_3 & -\mathbf{C}_4 \end{pmatrix} \tag{1f. 1. 10}$$

を一般逆行列の1つとする．このとき，

$$\inf_{\mathbf{B}'\mathbf{X}=\mathbf{U}} \mathbf{X}'\mathbf{A}\mathbf{X} = \mathbf{U}'\mathbf{C}_4\mathbf{U}. \tag{1f. 1. 11}$$

1f. 2　固有値と固有ベクトルに関連する結果

$\mathbf{A}$ を $m \times m$ の対称行列，$\lambda_1 \geqslant \cdots \geqslant \lambda_m$ をその固有値，$\mathbf{P}_1, \cdots, \mathbf{P}_m$ を対応する固有ベクトルとする．（1c. 3. 5）で，

$$\mathbf{A} = \lambda_1\mathbf{P}_1\mathbf{P}_1' + \cdots + \lambda_m\mathbf{P}_m\mathbf{P}_m'$$
$$\mathbf{I} = \mathbf{P}_1\mathbf{P}_1' + \cdots + \mathbf{P}_m\mathbf{P}_m'$$

であることを示した．次の諸命題が成立する．

（i）
$$\sup_{\mathbf{X}} \frac{\mathbf{X}'\mathbf{A}\mathbf{X}}{\mathbf{X}'\mathbf{X}} = \lambda_1, \qquad \inf_{\mathbf{X}} \frac{\mathbf{X}'\mathbf{A}\mathbf{X}}{\mathbf{X}'\mathbf{X}} = \lambda_m. \tag{1f. 2. 1}$$

任意の $\mathbf{X}$ は $C_1\mathbf{P}_1 + \cdots + C_m\mathbf{P}_m$ と書けることに注目せよ．すると，

$$\frac{\mathbf{X}'\mathbf{A}\mathbf{X}}{\mathbf{X}'\mathbf{X}} = \frac{C_1^2\lambda_1 + \cdots + C_m^2\lambda_m}{C_1^2 + \cdots + C_m^2}. \tag{1f. 2. 2}$$

ベクトル $(C_1, \cdots, C_m)$ に関するこの sup と inf は，明らかに λ_1 と λ_m である．sup は $\mathbf{X} = \mathbf{P}_1$ のとき，inf は $\mathbf{X} = \mathbf{P}_m$ のときに得られる．

（ii）
$$\sup_{\substack{\mathbf{P}_i'\mathbf{X}=0 \\ i=1,\cdots,k}} \frac{\mathbf{X}'\mathbf{A}\mathbf{X}}{\mathbf{X}'\mathbf{X}} = \lambda_{k+1}, \qquad \inf_{\substack{\mathbf{P}_i'\mathbf{X}=0 \\ i=1,\cdots,k}} \frac{\mathbf{X}'\mathbf{A}\mathbf{X}}{\mathbf{X}'\mathbf{X}} = \lambda_m. \tag{1f. 2. 3}$$

与えられた条件の下では，$\mathbf{X} = C_{k+1}\mathbf{P}_{k+1} + \cdots + C_m\mathbf{P}_m$．したがって，（i）で用いたと同じ議論で（1f. 2. 3）を得る．

（iii）$\mathbf{B}$ を $m \times k$ 行列とする．このとき，

$$\inf_{\mathbf{B}} \sup_{\mathbf{B}'\mathbf{X}=0} \frac{\mathbf{X}'\mathbf{A}\mathbf{X}}{\mathbf{X}'\mathbf{X}} = \lambda_{k+1} \tag{1f. 2. 4}$$

$$\sup_{\mathbf{B}} \inf_{\mathbf{B}'\mathbf{X}=0} \frac{\mathbf{X}'\mathbf{A}\mathbf{X}}{\mathbf{X}'\mathbf{X}} = \lambda_{m-k}. \tag{1f. 2. 5}$$

この値 λ_{k+1} は，（1f. 2. 4）で $\mathbf{B}$ が $\mathbf{P}_1, \cdots, \mathbf{P}_k$ を列とする行列であるときに得られる．また，値 λ_{m-k} は，$\mathbf{B}$ が $\mathbf{P}_{m-k+1}, \cdots, \mathbf{P}_m$ を列とする行列であるときに得られる．

（i）のときと同様に，$\mathbf{X} = C_1\mathbf{P}_1 + \cdots + C_m\mathbf{P}_m = \mathbf{P}\mathbf{C}$ とおく．ここで，$\mathbf{P}$ は $\mathbf{P}_i$ を第 i 列とする $m \times m$ 行列である．このとき，制約式 $\mathbf{B}'\mathbf{X} = \mathbf{0} \Leftrightarrow \mathbf{B}'\mathbf{P}\mathbf{C} = \mathbf{0}$ となり，$\mathbf{G} = \mathbf{P}'\mathbf{B}$ とおくと，$\mathbf{G}'\mathbf{C} = \mathbf{0}$ となる．すると，（1f. 2. 4）の問題は

$$\inf_{\mathbf{G}} \sup_{\mathbf{G}'\mathbf{C}=0} \frac{C_1^2\lambda_1 + \cdots + C_m^2\lambda_m}{C_1^2 + \cdots + C_m^2} \tag{1f. 2. 6}$$

を求めることに帰着する．ところで，$\mathbf{G}'\mathbf{C} = 0$，$C_{k+2} = \cdots = C_m = 0$ となるような $\mathbf{C}$ の部分集合を選ぶと，

$$\sup_{\mathbf{G}'\mathbf{C}=0} \frac{\sum_1^m C_i^2\lambda_i}{\sum_1^m C_i^2} \geqslant \sup_{\mathbf{G}'\mathbf{C}=0} \frac{\sum_1^{k+1} C_i^2\lambda_i}{\sum_1^{k+1} C_i^2} \geqslant \lambda_{k+1}. \tag{1f. 2. 7}$$

$\mathbf{C}$ のこのような部分集合は空集合ではない．（1f. 2. 7）の結果は $\mathbf{G}$ とは独立である．したがって，（1f. 2. 6）$\geqslant \lambda_{k+1}$．一方，（ii）によって，$\mathbf{P}_1, \cdots, \mathbf{P}_k$ をその列とする行列 $\mathbf{B}$ と，値 λ_{k+1} をとるような $\mathbf{X}$ が存在するから，（1f. 2. 4）は証明できる．同様に（1f. 2. 5）も証明される．

（iv）$\mathbf{X}_1, \cdots, \mathbf{X}_k$ を k 個の互いに直交するベクトルとする．このとき，

$$\sup_{\mathbf{X}_1,\cdots,\mathbf{X}_k} \sum_1^k \frac{\mathbf{X}_i'\mathbf{A}\mathbf{X}_i}{\mathbf{X}_i'\mathbf{X}_i} = \sum_1^k \lambda_i \tag{1f. 2. 8}$$

で，上限は $\mathbf{X}_i \propto \mathbf{P}_i$，$i = 1, \cdots, k$ のときに得られる．

$\mathbf{X}_i = C_{i1}\mathbf{P}_1 + \cdots + C_{im}\mathbf{P}_m$ は，付加条件 $\mathbf{X}_i'\mathbf{X}_i = 1$ をみたすように選ぶ（一般性を失うことなく）．このとき，

$$\sum \frac{\mathbf{X}_i'\mathbf{A}\mathbf{X}_i}{\mathbf{X}_i'\mathbf{X}_i} = \lambda_1(C_{11}^2 + \cdots + C_{k1}^2) + \cdots + \lambda_m(C_{1m}^2 + \cdots + C_{km}^2) \tag{1f. 2. 9}$$

と書け，各 λ_i の係数は $\leqslant 1$ で，係数の和は k になる．（1f. 2. 9）の左辺を最大にするための係数の最適な選択は，$\lambda_1, \cdots, \lambda_k$ の係数を1とし，残りの係数を0とすることであって，これは，$\mathbf{X}_i = \mathbf{P}_i$，$i = 1, \cdots, k$ に選べば可能である．

（v）$\mathbf{A}$ は階数 r の $m \times n$ 行列で，その特異値分解は $\mathbf{A} = \lambda_1\mathbf{P}_1\mathbf{Q}_1^* + \cdots + \lambda_r\mathbf{P}_r\mathbf{Q}_r^*$ とする（1c. 3 の(v)参照）．さらに $\mathscr{B}$ を階数 k（$<r$）の $m \times n$ 行列の族とする．このとき，

$$\inf_{\mathbf{B}\in\mathscr{B}} \|\mathbf{A} - \mathbf{B}\| = (\lambda_{k+1}^2 + \cdots + \lambda_r^2)^{1/2} \tag{1f. 2. 10}$$

で，その下限は，次のときに得られる：

$$\mathbf{B} = \lambda_1 \mathbf{P}_1 \mathbf{Q}_1^* + \cdots + \lambda_k \mathbf{P}_k \mathbf{Q}_k^*. \qquad (1f.2.11)$$

[ここで，記号 $\|\mathbf{A}\|$ は行列 $\mathbf{A} = (a_{ij})$ のユークリッド・ノルムで $(\sum\sum a_{ij}^2)^{1/2}$ で定義される.]

$\mathbf{C}$ を階数 k の $m \times k$ ユニタリ行列，$\mathbf{D}$ を階数 k の $k \times n$ 行列として，階数分解（1b.2,(iv) 参照）$\mathbf{B} = \mathbf{CD}$ を用いる．$\mathbf{C}$ を与えると，$\|\mathbf{A} - \mathbf{CD}\|$ の下限は，$\mathbf{D} = \mathbf{C}^*\mathbf{A}$ のとき，つまり，$\mathbf{CD}$ の第 i 列が $\mathbf{A}$ の第 i 列の $\mathscr{M}(\mathbf{C})$ 上への射影であるときに達せられる．このとき，

$$\|\mathbf{A} - \mathbf{CC}^*\mathbf{A}\|^2 = \operatorname{tr} \mathbf{AA}^* - \operatorname{tr} \mathbf{C}^*\mathbf{AA}^*\mathbf{C} \qquad (1f.2.12)$$

となる．(1f.2.12) を $\mathbf{C}$ について最小にするには，その右辺の第2項を最大にしなければならない．一方，(1f.2.8) から，

$$\sup_{\mathbf{C}} \operatorname{tr} \mathbf{C}^*\mathbf{AA}^*\mathbf{C} = \lambda_1^2 + \cdots + \lambda_k^2$$

で，$\lambda_1, \cdots, \lambda_k$ は $\mathbf{A}$ の特異値（$\lambda_1^2, \cdots, \lambda_k^2$ は $\mathbf{AA}^*$ の固有値）である．この最大は $\mathbf{C}$ の第 i 列 $\mathbf{C}_i$ が $\mathbf{AA}^*$ の第 i 固有ベクトルであるとき，つまり，$\mathbf{C}_i = \mathbf{P}_i$ のときに得られる．このとき，

$$\begin{aligned}\inf_{\mathbf{B} \in \mathscr{B}} \|\mathbf{A} - \mathbf{B}\|^2 &= \inf_{\mathbf{C}} \|\mathbf{A} - \mathbf{CC}^*\mathbf{A}\|^2 \\ &= \operatorname{tr} \mathbf{AA}^* - (\lambda_1^2 + \cdots + \lambda_k^2) = \lambda_{k+1}^2 + \cdots + \lambda_r^2\end{aligned}$$

となり，この下限は，$\mathbf{C}_i = \mathbf{P}_i$ のとき，つまり

$$\mathbf{B} = \mathbf{CC}^*\mathbf{A} = \lambda_1 \mathbf{P}_1 \mathbf{Q}_1^* + \cdots + \lambda_k \mathbf{P}_k \mathbf{Q}_k^*$$

のときに得られる．

この (1f.2.10) と (1f.2.11) の結果は，$\mathscr{B}$ を階数が k または k よりも小さい $m \times n$ 行列の族におきかえても成立することに注意せよ．

（vi）ストルムの分離定理　$r = 1, \cdots, m$ にたいして，行列の系列 $\mathbf{A}_r = (a_{ij})$, $i, j = 1, \cdots, r$ を考え，$\mathbf{A}_r$ の固有値を $\lambda_1(\mathbf{A}_r) \geqslant \cdots \geqslant \lambda_r(\mathbf{A}_r)$ であらわす．このとき，

$$\lambda_{k+1}(\mathbf{A}_{i+1}) \leqslant \lambda_k(\mathbf{A}_i) \leqslant \lambda_k(\mathbf{A}_{i+1}). \qquad (1f.2.13)$$

(1f.2.13) の結果は，(1f.2.1)，(1f.2.3) で与えられた行列の固有値の解釈を用いると，すぐに得られる．(vi) の系として，次の定理を得る．

（vii）ポアンカレの分離定理　$\mathbf{B}$ を $m \times k$ 行列で，$\mathbf{B}'\mathbf{B} = \mathbf{I}_k$（次数 k の単位行列）とする．このとき，

$$\lambda_i(\mathbf{B}'\mathbf{AB}) \leqslant \lambda_i(\mathbf{A}), \qquad i = 1, \cdots, k$$

$$\lambda_{k-j}(\mathbf{B}'\mathbf{AB}) \geqslant \lambda_{m-j}(\mathbf{A}), \qquad j = 0, 1, \cdots, k-1. \qquad (1f.2.14)$$

（viii）$\mathbf{A}$ を次数 m の p. d. 行列，$\mathbf{B}$ を階数 k の $m \times k$ 行列で $\mathbf{B}'\mathbf{B}$ の対角要素は所与の数 $g_1, \cdots, g_k$ であるとする．このとき，

$$\max_{\mathbf{B}} |\mathbf{B}'\mathbf{AB}| = g_1 \cdots g_k \lambda_1 \cdots \lambda_k. \qquad (1f.2.15)$$

$\mathbf{C}'\mathbf{C}$ の対角要素はすべて1とすると，

$$\max_{\mathbf{B}} |\mathbf{B}'\mathbf{AB}| = g_1 \cdots g_k \max_{\mathbf{C}} |\mathbf{C}'\mathbf{AC}|$$

となることに注目せよ．

さらに，$\gamma_1, \cdots, \gamma_k$ を $\mathbf{C}'\mathbf{C}$ の固有値とする．$\gamma_1 + \cdots + \gamma_k = k$ であるから，$\gamma_1 \cdots \gamma_k \leqslant 1$ であることがわかる．行列 $\mathbf{G}$ を，その第 i 列が $\mathbf{C}'\mathbf{C}$ の第 i 固有ベクトルの $\gamma_i^{-1/2}$ 倍になるように選ぶ．すると，次を得る．

（a）$\mathbf{D} = \mathbf{CG}$ とすると，$\mathbf{D}'\mathbf{D} = \mathbf{I}$

（b）$|\mathbf{D}'\mathbf{AD}| = |\mathbf{G}'\mathbf{C}'\mathbf{ACG}| = |\mathbf{C}'\mathbf{AC}||\mathbf{G}'\mathbf{G}|$

$$= |\mathbf{C}'\mathbf{AC}|(\gamma_1 \cdots \gamma_k)^{-1} \geqslant |\mathbf{C}'\mathbf{AC}|.$$

したがって，

$$\max_{\mathbf{C}} |\mathbf{C}'\mathbf{AC}| = \max_{\mathbf{D}'\mathbf{D}=\mathbf{I}} |\mathbf{D}'\mathbf{AD}|.$$

ここで，(1f.2.14) を適用すると，

$$|\mathbf{D}'\mathbf{AD}| = \prod_{i=1}^{k} \lambda_i(\mathbf{D}'\mathbf{AD}) \leqslant \prod_{i=1}^{k} \lambda_i(\mathbf{A})$$

を得る．ただし，$\mathbf{D}$ の列を $\mathbf{A}$ の最初の k 個の固有ベクトルにとると，等号が得られ，それは (1f.2.15) を証明することになる［問題20—22も参照せよ］.

$\mathbf{B}$ の列が直交すると仮定するのでなければ，

$$\min_{\mathbf{B}} |\mathbf{B}'\mathbf{AB}| = g_1 \cdots g_k \lambda_{m-k+1} \cdots \lambda_m$$

という結果は正しくないことに注意せよ．

1f.3　対角和最小問題

$\mathscr{C}_1$ は $m \times n$ 行列 $\mathbf{A}$ の族で，$\mathbf{X}, \mathbf{V}_i$ および p_i を所与の行列とスカラーとするとき，行列 $\mathbf{A}$ は $\mathbf{AX} = \mathbf{0}$，および $\operatorname{tr} \mathbf{AV}_i' = p_i$ を満足するとし，$\mathscr{C}_2$ を $\mathscr{C}_1$ と同じ族であるが，その $\mathbf{A}$ と $\mathbf{V}_i$ は対称な正方行列であるとする．さらに，$\mathbf{P}$ を $\mathscr{M}(\mathbf{X})$ 上への直交射影作用素，$\mathbf{Q} = \mathbf{I} - \mathbf{P}$ とし，その内積の定義は $(\mathbf{x}, \mathbf{y}) = \mathbf{y}^*\mathbf{x}$ とする．

（i）$\mathbf{A} \in \mathscr{C}_1$ のとき，$\operatorname{Min} \operatorname{tr} \mathbf{AA}'$ は

$$\mathbf{A}_* = \sum \lambda_i \mathbf{V}_i \mathbf{Q} \qquad (1f.3.1)$$

で達せられる．ただし，$\lambda_1, \cdots, \lambda_k$ は

$$\sum_{i=1}^{k} \lambda_i \operatorname{tr} \mathbf{V}_i \mathbf{Q} \mathbf{V}_j' = p_j, \qquad j = 1, \cdots, k \qquad (1f.3.2)$$

の解である．

$\mathbf{A}_*\mathbf{X} = \mathbf{0}$ であることに注意する．ここで，$\mathbf{A}_*$ のかわりに $\mathbf{A}_* + \mathbf{D}$ を考えると，$\operatorname{tr} \mathbf{DV}_i'$

$=0$，および $\mathbf{DX}=0 \Rightarrow \mathbf{QD}'=\mathbf{D}'$ となる．

$$\operatorname{tr}\mathbf{A}_*\mathbf{D}' = \sum \lambda_i \operatorname{tr}\mathbf{V}_i\mathbf{QD}' = \sum \lambda_i \operatorname{tr}\mathbf{V}_i\mathbf{D}' = 0$$

となることを確かめよ．このとき，

$$\operatorname{tr}(\mathbf{A}_*+\mathbf{D})(\mathbf{A}_*+\mathbf{D})' = \operatorname{tr}\mathbf{A}_*\mathbf{A}_*' + \operatorname{tr}\mathbf{DD}' \geqslant \operatorname{tr}\mathbf{A}_*\mathbf{A}_*'$$

となり，求める結果が証明された．式 (1f. 3. 2) は解をもつことに注意せよ．

（ii）$\mathbf{A}\in\mathscr{C}_2$ のときの $\operatorname{Min}\operatorname{tr}\mathbf{AA}$ は

$$\mathbf{A}_* = \sum \lambda_i \mathbf{QV}_i\mathbf{Q} \tag{1f. 3. 3}$$

で得られる．ただし，$\lambda_1,\cdots,\lambda_k$ は，

$$\sum \lambda_i \operatorname{tr}\mathbf{QV}_i\mathbf{QV}_j = p_j, \qquad j=1,\cdots,k \tag{1f. 3. 4}$$

の解である．

この結果は（i）と同様に証明できる．

（iii）$\mathbf{V}$ を p. d. 行列とする．このとき，$\mathbf{A}\in\mathscr{C}_2$ なら，$\min\operatorname{tr}\mathbf{AVAV}$ は，

$$\mathbf{A}_* = \sum \lambda_i \mathbf{Q}_V'\mathbf{V}^{-1}\mathbf{V}_i\mathbf{V}^{-1}\mathbf{Q}_V \tag{1f. 3. 5}$$

で得られる．ただし，$\lambda_1,\cdots,\lambda_k$ は

$$\sum \lambda_i \operatorname{tr}\mathbf{Q}_V'\mathbf{V}^{-1}\mathbf{V}_i\mathbf{V}^{-1}\mathbf{Q}_V\mathbf{V}_j = p_j, \qquad j=1,\cdots,k \tag{1f. 3. 6}$$

の解で，$\mathbf{Q}_V=\mathbf{I}-\mathbf{P}_V$，　$\mathbf{P}_V=\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^-\mathbf{X}'\mathbf{V}^{-1}$ は $\mathscr{M}(\mathbf{X})$ 上への直交射影作用素，その内積の定義は $(\mathbf{x},\mathbf{y})=\mathbf{y}^*\mathbf{V}^{-1}\mathbf{x}$ である．

この結果は，行列 $\mathbf{B}=\mathbf{V}^{\frac{1}{2}}\mathbf{AV}^{\frac{1}{2}}$ を考え，(ii) の結果を適用すると，証明できる．

$\mathscr{C}_3$ は対称行列 $\mathbf{A}$ の族で，$\mathbf{X}'\mathbf{AX}=0$，および $\operatorname{tr}\mathbf{AV}_i=p_i$，$i=1,\cdots,k$ をみたすものとする．ただし，$\mathbf{V}_i$ は対称とする．

（iv）$\mathbf{A}\in\mathscr{C}_3$ のとき，$\operatorname{Min}\operatorname{tr}\mathbf{AA}$ は，

$$\mathbf{A}_* = \sum \lambda_i(\mathbf{V}_i - \mathbf{PV}_i\mathbf{P}) \tag{1f. 3. 7}$$

で得られる．ただし $\lambda_1,\cdots,\lambda_k$ は

$$\sum \lambda_i \operatorname{tr}(\mathbf{V}_i-\mathbf{PV}_i\mathbf{P})\mathbf{V}_j = p_j, \qquad j=1,\cdots,k \tag{1f. 3. 8}$$

の解である．

ここで，$\mathbf{X}'\mathbf{AX}=0$ の一般解は $\mathbf{T}-\mathbf{PTP}$ の形をとること，および，$\mathbf{A}_*$ のかわりに $\mathbf{A}_*+\mathbf{T}-\mathbf{PTP}$ と書くことができ，$\operatorname{tr}(\mathbf{T}-\mathbf{PTP})\mathbf{V}_i=0$ という条件が満足されることに注意する．すると，

$$\operatorname{tr}(\mathbf{V}_i-\mathbf{PV}_i\mathbf{P})(\mathbf{T}-\mathbf{PTP}) = \operatorname{tr}\mathbf{V}_i(\mathbf{T}-\mathbf{PTP}) = 0.$$

このとき，

$$\begin{aligned}\operatorname{tr}(\mathbf{A}_*+\mathbf{T}-\mathbf{PTP})(\mathbf{A}_*+\mathbf{T}-\mathbf{PTP}) &= \operatorname{tr}\mathbf{A}_*\mathbf{A}_* + \operatorname{tr}(\mathbf{T}-\mathbf{PTP})(\mathbf{T}-\mathbf{PTP}) \\ &\geqslant \operatorname{tr}\mathbf{A}_*\mathbf{A}_*\end{aligned}$$

となり，求める結果が証明された．

（v）$\mathbf{V}$ を p. d. 行列とする．ここで，$\mathbf{A}\in\mathscr{C}_3$ のときの $\min\operatorname{tr}\mathbf{AVAV}$ は

$$\mathbf{A}_* = \sum \lambda_i \mathbf{V}^{-1}(\mathbf{V}_i - \mathbf{P}_V\mathbf{V}_i\mathbf{P}_V')\mathbf{V}^{-1} \tag{1f. 3. 9}$$

で得られる．ただし，$\lambda_1,\cdots,\lambda_k$ は

$$\sum \lambda_i \operatorname{tr}\mathbf{V}^{-1}(\mathbf{V}_i - \mathbf{P}_V\mathbf{V}_i\mathbf{P}_V')\mathbf{V}^{-1}\mathbf{V}_j = p_j, \qquad j=1,\cdots,k \tag{1f. 3. 10}$$

の解である．

この結果も，変換 $\mathbf{V}^{\frac{1}{2}}\mathbf{AV}^{\frac{1}{2}}=\mathbf{B}$ を施した後に，(iv) の結果を適用すると得られる．

本章で論議した特殊な行列の定義を表1にまとめる．

表 1

行列 A の型	定　　義
対称	$\mathbf{A}=\mathbf{A}'$（$\mathbf{A}'$ は転置行列）
エルミート	$\mathbf{A}=\mathbf{A}^*$（$\mathbf{A}^*$ は共役転置行列）
ベキ等	$\mathbf{A}^2=\mathbf{A}$
3ベキ等	$\mathbf{A}^3=\mathbf{A}$
ベキ零	$\mathbf{A}^k=0$，ある整数 k について
p. d.（正定符号）	$\mathbf{X}^*\mathbf{AX}>0$，すべての非零 $\mathbf{X}$ について
p. s. d.（半正定符号）	$\mathbf{X}^*\mathbf{AX}\geqslant 0$，すべての $\mathbf{X}$ について，かつ，$\mathbf{X}^*\mathbf{AX}=0$，ある非零 $\mathbf{X}$ について
n. n. d.（非負定符号）	$\mathbf{X}^*\mathbf{AX}\geqslant 0$，すべての $\mathbf{X}$ について
正規	$\mathbf{AA}^*=\mathbf{A}^*\mathbf{A}$
半単純	$\mathbf{A}=\mathbf{P}\Lambda\mathbf{P}^{-1}$，ある $\mathbf{P}$ と対角行列 Λ について
直交	$\mathbf{AA}'=\mathbf{A}'\mathbf{A}=\mathbf{I}$
ユニタリ	$\mathbf{AA}^*=\mathbf{A}^*\mathbf{A}=\mathbf{I}$
g 逆行列（一般逆行列）	$\mathbf{A}^-$
反射型一般逆行列	A_r^-
N ノルム最小型一般逆行列	$A_{m(\mathrm{N})}^-$　[(1c. 5. 1) を見よ]
M 最小2乗型一般逆行列	$A_{l(\mathrm{M})}^-$　[(1c. 5. 3) を見よ]
N ノルム最小・M 最小2乗型一般逆行列	A_{MN}^+　[(1c. 5. 5) を見よ]

補足と問題

1　行列の固有値．

1.1　$m\times m$ 行列

$$(a_{ij}) = \begin{pmatrix} a & b & \cdots & b \\ b & a & \cdots & b \\ \cdots & \cdots & \cdots & \cdots \\ b & b & \cdots & a \end{pmatrix}$$

について，次を証明せよ．

（i）$|a_{ij}| = (a-b)^{n-1}(a+\overline{n-1}b)$.

（ii）逆行列もまた同じ形になり，

$$a^{ii} = \frac{a+\overline{n-2}b}{(a+\overline{n-1}b)(a-b)}, \qquad a^{ij} = \frac{-b}{(a+\overline{n-1}b)(a-b)}, \quad i \neq j.$$

（iii）この行列の固有値は，$a+\overline{n-1}b$ でその固有ベクトルは $(1,1,\cdots,1)$ のものと，$n-1$ 重根をとる $a-b$ でその固有ベクトルは $(1,1,\cdots,1)$ に直交する（任意の）$n-1$ 個のベクトルである．

1.2　巡回行列

$$(a_{ij}) = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ a_n & a_1 & \cdots & a_{n-1} \\ & \cdots\cdots\cdots\cdots & & \\ a_2 & a_3 & \cdots & a_1 \end{pmatrix}$$

について，次の諸結果を証明せよ．

（i）$|a_{ij}| = \prod_{i=1}^{n}[a_1 + a_2\omega_i + a_3\omega_i^2 + \cdots + a_n\omega_i^{n-1}]$，ただし，$\omega_i$ は1の i 乗根である．

（ii）この固有値は，$\lambda_i = a_1 + a_2\omega_i + \cdots + a_n\omega_i^{n-1}$，$i = 1,\cdots,n$ で，対応する固有ベクトルは，$(1,\omega_i,\cdots,\omega_i^{n-1})$，$i = 1,\cdots,n$ である．

1.3　問題 **1.1** の結果を，

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} & \cdots & \mathbf{B} \\ \mathbf{B} & \mathbf{A} & \cdots & \mathbf{B} \\ & \cdots\cdots\cdots\cdots & & \\ \mathbf{B} & \mathbf{B} & \cdots & \mathbf{A} \end{pmatrix}$$

の形の行列に拡張せよ．ただし，$\mathbf{A}, \mathbf{B}$ はそれら自身，次数 k の正方行列である．特に，その行列式の値が $|\mathbf{A}-\mathbf{B}|^{n-1}|\mathbf{A}+\overline{n-1}\mathbf{B}|$ になることを示せ．

1.4　$\mathbf{U}$ を $m \times n$ 行列とする．$|\mathbf{UU}' - \lambda\mathbf{I}| = 0$ と $|\mathbf{U}'\mathbf{U} - \lambda\mathbf{I}| = 0$ の零でない根は一致する．[$|\mathbf{UU}' - \lambda\mathbf{I}| = 0$ の零でない根を λ^* とすると，$\mathbf{UU}'\mathbf{L} = \lambda^*\mathbf{L}$ をみたすベクトル $\mathbf{L}$ が存在する．この両辺に $\mathbf{U}'$ をかけると，$\mathbf{U}'\mathbf{UU}'\mathbf{L} = \lambda^*\mathbf{U}'\mathbf{L}$ となり，$(\mathbf{U}'\mathbf{U} - \lambda^*\mathbf{I})\mathbf{U}'\mathbf{L} = \mathbf{0}$ と書ける．したがって，λ^* は $\mathbf{U}'\mathbf{U}$ の固有値でもある．]

1.5　$\mathbf{UU}'$ または $\mathbf{U}'\mathbf{U}$ の零でない固有値を $\lambda_1,\cdots,\lambda_r$ とするとき，$|\mathbf{UU}' - \mathbf{I}| = (\lambda_1 - 1)\cdots(\lambda_r - 1)(0-1)^{m-r}$ および $|\mathbf{U}'\mathbf{U} - \mathbf{I}| = (\lambda_1 - 1)\cdots(\lambda_r - 1)(0-1)^{n-r}$ を証明せよ．

1.6　$\mathbf{U}$ を階数 r の $m \times n$ 行列とする．長さ1のベクトル $\mathbf{C}_1$ と $\mathbf{C}_2$ をいろいろに変えたときの $[\mathbf{C}_1'\mathbf{UC}_2]^2$ の停留値は $\mathbf{UU}'$ の固有値である．

1.7　直交行列について，(a)　すべての固有値の絶対値は1である．(b)　λ が固有値なら $1/\lambda$ もまた固有値である．および，(c)　その行列次数が奇数なら，少なくとも1つの根は ± 1 である，ことを示せ．

1.8　$\lambda_1, \lambda_2, \cdots, \lambda_k$ を正方行列 $\mathbf{A}$ の固有値とすると，(a)　$\mathrm{Trace}\,\mathbf{A} = \sum \lambda_i$，(b)　$\mathrm{trace}\,\mathbf{A}'\mathbf{A} \geqq \sum \lambda_i^2$，および，(c)　$|\mathbf{A}| = \lambda_1 \cdots \lambda_k$ である．

1.9　$\mathbf{A}$ と $\mathbf{B}$ を次数 m の対称行列とする．$\mathbf{C}$ の i 番目の大きさの固有値を $\lambda_i(\mathbf{C})$ であらわすと，

$$\left.\begin{aligned} \lambda_i(\mathbf{A}+\mathbf{B}) &\leqslant \lambda_j(\mathbf{A}) + \lambda_k(\mathbf{B}) \\ \lambda_{m+i-1}(\mathbf{AB}) &\geqslant \lambda_{m-j+1}(\mathbf{A})\lambda_{m-k+1}(\mathbf{B}) \end{aligned}\right\} \quad j+k \leqslant i+1.$$

(Anderson and Dasgupta (1963))

2　$\mathbf{A}$ を階数 r の $n \times m$ 行列，$\lambda_1,\cdots,\lambda_r$ を $\mathbf{A}'\mathbf{A}$ の零でない固有値，$\mathbf{P}_1,\cdots,\mathbf{P}_r$ をその固有ベクトルとする．このとき，$\lambda_1^{-1}\mathbf{P}_1\mathbf{P}_1' + \cdots + \lambda_r^{-1}\mathbf{P}_r\mathbf{P}_r'$ は $\mathbf{A}'\mathbf{A}$ の g 逆行列である．

3　p.d. または p.s.d. 行列はグラム行列ともよび，$\mathbf{G}$ であらわす．次の結果，(そのいくつかはすでに証明したが) は重要な応用をもっている．

3.1　$\mathbf{M}$ は $k \times p$ で階数 r とすると，$\mathbf{MM}'$ は階数 r のグラム行列である．

3.2　次数 k で階数 r の行列 $\mathbf{G}$ は，階数 r の $k \times r$ 行列 $\mathbf{F}$ を用いて，$\mathbf{G} = \mathbf{FF}'$ と分解できる．$k \times s$ 行列 $\mathbf{M}$ を用いて $\mathbf{MM}' = \mathbf{FF}' = \mathbf{G}$ と書けるなら，$\mathbf{M} = (\mathbf{F} \vdots \mathbf{0})\mathbf{U}$ とするような直交行列 $\mathbf{U}$ が存在する．

3.3　$\mathbf{G}$ が正則なら，その主小行列はすべて正則である．

3.4　積 $\mathbf{MGM}'$ が存在するかぎり，任意の $\mathbf{M}$ について $\mathbf{MGM}'$ はグラム行列である．

3.5　$\mathbf{G}$ には一意のグラム平方根行列が存在する．

[非負の固有値 $\lambda_1,\cdots,\lambda_n$ の対角行列を Λ と書くとき，$\mathbf{G} = \mathbf{U}\Lambda\mathbf{U}'$ なる直交行列 $\mathbf{U}$ が存在する．$\mathbf{M} = \mathbf{U}\Lambda^{1/2}\mathbf{U}'$ は (a)　対称で，(b)　固有値 $\lambda_1^{1/2},\cdots,\lambda_n^{1/2}$ をもち，したがって，グラム行列であることを証明せよ．さらに，$\mathbf{MM} = \mathbf{U}\Lambda^{1/2}\mathbf{U}'\mathbf{U}\Lambda^{1/2}\mathbf{U}' = \mathbf{U}\Lambda\mathbf{U}' = \mathbf{G}$ であるから，$\mathbf{M}$ は求める平方根である．そのような平方根がほかにあっても，すべて，同じ固有値 $\lambda_1^{1/2},\cdots,\lambda_n^{1/2}$ と固有ベクトルをもつことになるから，これは確かに一意である．]

3.6　$\mathbf{T}$ を対角要素が非負の三角行列とするとき，$\mathbf{G} = \mathbf{TT}'$ なる分解は一意である．

[証明は，$\mathbf{G}$ とそのような行列 $\mathbf{T}$ との1：1対応を示すことによってできる．]

3.7　$\mathbf{G}_1$ と $\mathbf{G}_2$ が n.n.d. なら，$\mathbf{G}_1 + \mathbf{G}_2$ も n.n.d. で，$\mathscr{M}(\mathbf{G}_1) \subset \mathscr{M}(\mathbf{G}_1 + \mathbf{G}_2)$，$\mathscr{M}(\mathbf{G}_2) \subset \mathscr{M}(\mathbf{G}_1 + \mathbf{G}_2)$ である．

4　ラグランジュの定理　$\mathbf{S}$ を次数 n，階数 $r > 0$ の正方行列，$\mathbf{X}, \mathbf{Y}$ を $\mathbf{X}'\mathbf{SY} \neq 0$ なる列ベクトルとする．このとき，残差行列

$$\mathbf{S}_1 = \mathbf{S} - \frac{\mathbf{SYX}'\mathbf{S}}{\mathbf{X}'\mathbf{SY}}$$

の階数は正確に $r-1$ である．

より一般的に，$\mathbf{S}$ を階数 $r > 0$ の $n \times m$ 行列，$\mathbf{A}$ と $\mathbf{B}$ をそれぞれ $s \times n$ および $s \times m$ 行列で $s \leqslant r$，かつ $\mathbf{ASB}'$ は正則とすると，残差行列

$$\mathbf{S}_1 = \mathbf{S} - \mathbf{SB}'(\mathbf{ASB}')^{-1}\mathbf{AS}$$

の階数は，正確に $r-s$ となり，$\mathbf{S}$ がグラム行列なら $\mathbf{S}_1$ もグラム行列となる．

5　行列 (x_{ij})，$i = 1,\cdots,m$；$j = 1,\cdots,n$ を考え，

$$\sum_j x_{ij} = n\bar{x}_i, \qquad S_{tu} = \sum_j (x_{tj} - \bar{x}_t)(x_{uj} - \bar{x}_u)$$

と定義する．このとき，$m \times m$ 行列 (S_{tu}) は次の性質をもつ．

（i）$(x_{ij} - \bar{x}_i)$ の階数が m のとき正定符号である，

（ii）$(x_{ij} - \bar{x}_i)$ の階数が m より小さい，すなわち

$$\sum p_i(x_{ij} - \bar{x}_i) = 0, \qquad j = 1,\cdots,n$$

なる定数 $p_1,\cdots,p_m$ が存在するとき，半正定符号である．

（iii）$n < m$ のとき必ず，半正定符号である．

6　行列の標準化.

6.1　$\mathbf{A}$を次数mの正方行列とするとき，$\mathbf{AC}$は三角行列で，その主対角線の下の要素はすべて零，対角要素は$\geqslant 0$とするような，直交行列$\mathbf{C}$が存在する.

6.2　極形式標準化　$|\mathbf{A}| \neq 0$ なら，$\mathbf{G}$をp. d.，$\mathbf{H}$を直交行列として $\mathbf{A} = \mathbf{GH}$ と書ける.［$\mathbf{G}$を$\mathbf{A}'\mathbf{A}$の一意のグラム平方根行列とせよ．$|\mathbf{A}| \neq 0$であるから$\mathbf{G}$はp. d.．このとき$\mathbf{H}$は$\mathbf{G}^{-1}\mathbf{A}$とする.］

6.3　$\mathbf{A}$を次数m，階数mの対称行列とする．三角行列$\mathbf{T}$が存在して，$\mathbf{A} = \mathbf{T}'\Delta^{-1}\mathbf{T}$ とできる．ただし，Δはその対角要素が$\mathbf{T}$と同じ対角行列である.

7　$\mathbf{A} = \mathbf{B} + \mathbf{C}$ で，$\mathbf{B}$はp. d.，$\mathbf{C}$は歪対称なら，$|\mathbf{A}| \geqslant |\mathbf{B}|$.

8　$\mathbf{A}$と$\mathbf{B}$を次数nの実p. d. 行列とすると，

（i）　$|\lambda\mathbf{A} + (1-\lambda)\mathbf{B}| \geqslant |\mathbf{A}|^{\lambda}|\mathbf{B}|^{1-\lambda}$,　$0 \leqslant \lambda \leqslant 1$,

（ii）　$|\mathbf{A} + \mathbf{B}|^{1/n} \geqslant |\mathbf{A}|^{1/n} + |\mathbf{B}|^{1/n}$.

9　$\mathbf{B}$をp. d.，$\mathbf{A} - \mathbf{B}$をn. n. d. とすると，

（i）　$|\mathbf{A} - \lambda\mathbf{B}| = 0$ の根はすべて$\lambda \geqslant 1$で，逆に，根がすべて$\lambda \geqslant 1$なら$\mathbf{A} - \mathbf{B}$はn. n. d. である.

（ii）　$|\mathbf{A}| \geqslant |\mathbf{B}|$.［(i) の結果として，$|\mathbf{A}|/|\mathbf{B}| = \lambda_1\lambda_2\cdots \geqslant 1$．この結果から，$\mathbf{C}$が，p. d. で$\mathbf{D}$がn. n. d. なら $|\mathbf{C}| \leqslant |\mathbf{C} + \mathbf{D}|$ が出る.］

（iii）　$\mathbf{B}^{-1} - \mathbf{A}^{-1}$はn. n. d. である.［これは，$|\mathbf{B}^{-1} - \mu\mathbf{A}^{-1}| = 0$ の根がすべて$\mu \geqslant 1$なら正しい．しかし，$|\mathbf{B}^{-1} - \mu\mathbf{A}^{-1}| = 0$ の根は $|\mathbf{A} - \mu\mathbf{B}| = 0$ の根と同じである．よって (i) から $\mu \geqslant 1$.］

（iv）　$\mathbf{A}_r$を$\mathbf{A}$の次数rの主小行列，$\mathbf{B}_r$をそれに対応する$\mathbf{B}$の小行列とすると，$\mathbf{A}_r - \mathbf{B}_r$はn. n. d. である.

10　$\mathbf{A}$を $m \times n$ 行列とする．$\|\mathbf{A} - \mathbf{B}\|$ が最小になるような階数rの $m \times n$ 行列$\mathbf{B}$を求めよ.

［ヒント：行列$\mathbf{AA}'$の最初のr個の固有ベクトルを考え，それらを$\mathbf{P}_1, \cdots, \mathbf{P}_r$とする．$\alpha_i$を$\mathbf{A}$の$i$番目の列ベクトルとすると，$\|\mathbf{A} - \mathbf{B}\|$ を最小とするような$\mathbf{B}_*$のi番目の列は$(\mathbf{P}_1'\alpha_i)\mathbf{P}_1 + \cdots + (\mathbf{P}_r'\alpha_i)\mathbf{P}_r$ となる．$\mathbf{B}$の基底をつくるr個の正規直交ベクトルをどのように選んでもユークリッド・ノルム $\|\mathbf{A} - \mathbf{B}\|$ は必ずより大きい値になることを示せ.］

11　(a_{ij}) を非負の要素から成る $m \times t$ 行列とし，

$$\sum_i a_{ij} = a_{\cdot j}, \qquad \sum_j a_{ij} = a_{i\cdot}, \qquad \sum_i a_{i\cdot} = \sum_j a_{\cdot j} = a_{\cdot\cdot}$$

とおく．このとき，次が成立する.

（i）　$mt \sum\sum a_{ij}a_{i\cdot}a_{\cdot j} \geqslant (a_{\cdot\cdot})^3$.

（ii）　$\sum\sum a_{ij}(p_i \sum_1^t a_{is}q_s)(q_j \sum_1^m a_{rj}p_r) \geqslant (\sum\sum a_{ij}p_iq_j)^3$,　ただし，$p_i, q_j$ は

$$\sum_1^m p_i = \sum_1^t q_j = 1, \qquad p_i, q_j > 0$$

なる任意の数とする.

（iii）　$m = t$, $a_{ij} = a_{ji}$, $p_i = q_i$ という特殊の場合には，

$$\sum\sum a_{ij}p_i'p_j' \geqslant \sum\sum a_{ij}p_ip_j.$$

ただし，$p_i'(\sum\sum a_{ij}p_ip_j) = \sum a_{is}p_s$ とする．（この不等式は，遺伝理論のなかで自然に出てくる．$A_1, \cdots, A_m$を1つの遺伝子座にある対立遺伝子とし，遺伝子型A_iA_jをもつ2倍体の相対生活力を$a_{ij} \geqslant 0$であらわす．p_iをある世代におけるA_iの〔相対〕度数とすると，無作為交配された母集団の平均生活力は $\sum\sum a_{ij}p_ip_j$ となる．A_iの度数p_iは $(\sum a_{is}p_s)/(\sum\sum a_{ij}p_ip_j)$ に変わり，次の世代の平均生活力は $\sum\sum a_{ij}p_i'p_j'$ となる．上の不等式は，世代が次へ進むにつれて，平均生活力は減少しえないことを主張している.）［Atkinson, Watterson and Moran (1960)］

12　$\mathbf{X}$を $k \times 1$ の列ベクトル，$\mathbf{A}$を次数kの正方対称行列で，全部の要素が非負であるとする．このとき，任意の整数nにたいして，

$$(\mathbf{X}'\mathbf{A}^n\mathbf{X})(\mathbf{X}'\mathbf{X})^{n-1} \geqslant (\mathbf{X}'\mathbf{AX})^n$$

で，等号は，$\mathbf{X}$が$\mathbf{A}$の固有ベクトルであるときにのみ成立する．この不等式は遺伝理論に現われる (Mulholland and Smith (1959)).

13　同時導関数を求める規則.

13.1　ベクトル変数$\mathbf{X}$のスカラー関数を $f(\mathbf{X})$ とし，$\mathbf{X}' = (X_1, \cdots, X_n)$ と書き，$\partial f/\partial X_1, \cdots, \partial f/\partial X_n$ を要素とする列ベクトルを $\partial f/\partial \mathbf{X}$ であらわす．同様に，行列 $(\partial^2 f/\partial X_i \partial X_j)$ として $\partial^2 f/\partial \mathbf{X}^2$ を定義する．次を証明せよ.

スカラー関数	ベクトル導関数	行列導関数
$f(\mathbf{X})$	$\partial f/\partial \mathbf{X}$	$\partial^2 f/\partial \mathbf{X}^2$
$\mathbf{M}'\mathbf{X}$	$\mathbf{M}$	$\mathbf{0}$
$\mathbf{X}'\mathbf{AZ}$	$\mathbf{AZ}$	$\mathbf{0}$
$\mathbf{X}'\mathbf{X}$	$2\mathbf{X}$	$2\mathbf{I}$
$\mathbf{X}'\mathbf{AX}$	$2\mathbf{AX}$	$2\mathbf{A}$ ($\mathbf{A}$が対称のとき)

13.2　$f(\mathbf{B})$を行列$\mathbf{B} = (b_{ij})$のスカラー関数とし，行列 $(\partial f(\mathbf{B})/\partial b_{ij})$ を $\partial f(\mathbf{B})/\partial \mathbf{B}$ であらわす．次を証明せよ.

（i）　$\dfrac{\partial}{\partial \mathbf{B}}(\mathbf{Y}'\mathbf{BZ}) = \mathbf{YZ}'$

（ii）　$\dfrac{\partial}{\partial \mathbf{B}}\mathrm{tr}\,\mathbf{B} = \mathbf{I}$

（iii）　$\dfrac{\partial}{\partial \mathbf{B}}\mathbf{X}'\mathbf{BX} = \mathbf{XX}'$,　$\mathbf{B}$が非対称のとき

$= 2\mathbf{XX}' - \mathrm{diag}\,\mathbf{XX}'$,　$\mathbf{B}$が対称のとき

（iv）　$\dfrac{\partial|\mathbf{B}|}{\partial \mathbf{B}} = |\mathbf{B}|(\mathbf{B}^{-1})'$,　$\mathbf{B}$が非対称のとき

$= |\mathbf{B}|(2\mathbf{B}^{-1} - \mathrm{diag}\,\mathbf{B}^{-1})$,　$\mathbf{B}$が対称のとき

（v）　$\dfrac{\partial\,\mathrm{tr}\,\mathbf{BC}}{\partial \mathbf{B}} = \mathbf{C}'$,　$\mathbf{B}$が非対称のとき

$= \mathbf{C} + \mathbf{C}' - \mathrm{diag}\,\mathbf{C}$,　$\mathbf{B}$が対称のとき

（vi）　$\dfrac{\partial\,\mathrm{tr}\,\mathbf{BCBD}}{\partial \mathbf{B}} = \mathbf{T}'$,　$\mathbf{B}$が非対称のとき

$= \mathbf{T} + \mathbf{T}' - \mathrm{diag}\,\mathbf{T}$,　$\mathbf{B}$が対称のとき

ただし，$\mathbf{T} = \mathbf{DBC} + \mathbf{CBD}$ で，$\mathrm{diag}\,\mathbf{T}$は$\mathbf{T}$と同じ対角要素をもつ対角行列.

これらの規則を $(\mathbf{Y} - \mathbf{AX})'(\mathbf{Y} - \mathbf{AX})$ を$\mathbf{X}$について最小にするのに適用せよ．ここで，

$\mathbf{A}$は次数 $m \times n$ 行列，$\mathbf{Y}$は $m \times 1$ 行列で所与である．最小にする方程式は $\mathbf{A}'\mathbf{A}\bar{\mathbf{X}} = \mathbf{A}'\mathbf{Y}$ であることを示せ．

14　$\mathbf{R}$を相関行列 (n. n. d. で，対角要素は 1)，$\mathbf{D}^2$は対角行列で $0 \leqslant d_i^2 \leqslant 1$ とする．$\inf_{\mathbf{D}} \operatorname{rank}(\mathbf{R} - \mathbf{D}^2) \geqslant s$ を証明せよ．ただし，sは $\mathbf{R}$の固有値で，1より大きいものの数である (Guttman (1954))．

15　$\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_t$ を対称正方行列とする．$\mathbf{C}\mathbf{A}_1\mathbf{C}', \cdots, \mathbf{C}\mathbf{A}_t\mathbf{C}'$ をすべて対角型にするような直交行列$\mathbf{C}$が存在するための必要十分条件は，すべての i と j について $\mathbf{A}_i\mathbf{A}_j$ が対称である．すなわち，$\mathbf{A}_i$ と $\mathbf{A}_j$ が交換可能であることである．

16　行列$\mathbf{A}$は，$\mathbf{A}^2 = \mathbf{A}$ のとき，ベキ等と言う．対称ベキ等行列$\mathbf{A}$についての次の諸結果を証明せよ．

16.1　$\mathbf{A}$の固有値はどれも0か1である．

16.2　$|\mathbf{A}| \neq 0 \Rightarrow \mathbf{A} = \mathbf{I}$ (単位行列)．

16.3　$\mathbf{A}$は非負定符号で，$\mathbf{L}_1, \cdots, \mathbf{L}_r$ を正規直交ベクトル，$r = \operatorname{rank}\mathbf{A}$ とすると，$\mathbf{A} = \mathbf{L}_1\mathbf{L}_1' + \cdots + \mathbf{L}_r\mathbf{L}_r'$ という表現ができる．

16.4　$\mathbf{A}$の第 i 対角要素が零なら，i 行，i 列のすべての要素は零である．

16.5　$\operatorname{trace}\mathbf{A} = \operatorname{rank}\mathbf{A}$.

17　$\mathbf{A}_1, \cdots, \mathbf{A}_t$ は対称とする．$\sum \mathbf{A}_i$ がベキ等で，$i \neq j$ にたいして $\mathbf{A}_i\mathbf{A}_j = 0$ なら，各 $\mathbf{A}_i$ はベキ等で，したがって $\operatorname{rank}(\sum \mathbf{A}_i) = \sum \operatorname{rank}\mathbf{A}_i$ となる．

18　非対称行列のスペクトル分解の記号 (1c. 3. 14)

$$\mathbf{A} = \delta_1\mathbf{P}_1\mathbf{Q}_1' + \cdots + \delta_m\mathbf{P}_m\mathbf{Q}_m'$$

を用いて，$\mathbf{A}$は

$$\mathbf{A} = \delta_1\mathbf{A}_1 + \cdots + \delta_m\mathbf{A}_m$$

と書けることを示せ．ただし，$\sum \mathbf{A}_i = \mathbf{I}$，$\mathbf{A}_i$はベキ等で，$\mathbf{A}_i\mathbf{A}_j = 0$，$i \neq j$ である．さらに，$L_j(z)$ をラグランジュの補間多項式，

$$L_j(z) = \prod_{\substack{i=1 \\ i \neq j}}^{m} \frac{z - \delta_i}{\delta_j - \delta_i}$$

とすると，$\mathbf{A}_j = L_j(\mathbf{A})$ である．[ヒント：$\mathbf{A}_i = \mathbf{P}_i\mathbf{Q}_i'$ に選べ．]

19　**有限次元ベクトル空間 $\mathscr{S}$ の2つの内積間の関係**　$\mathbf{X}, \mathbf{Y}$ を2つのベクトルとし，2つの可能な内積を $(\mathbf{X}, \mathbf{Y})_1$ と $(\mathbf{X}, \mathbf{Y})_2$ で定義する．

19.1

$$\frac{(\mathbf{P}_1, \mathbf{P}_1)_1}{(\mathbf{P}_1, \mathbf{P}_1)_2} = \sup_{\mathbf{X}} \frac{(\mathbf{X}, \mathbf{X})_1}{(\mathbf{X}, \mathbf{X})_2}$$

を満足するようなベクトル$\mathbf{P}_1$が存在するための必要十分条件は，任意の$\mathbf{U}$にたいして，$(\mathbf{P}_1, \mathbf{U})_1 = 0 \Leftrightarrow (\mathbf{P}_1, \mathbf{U})_2 = 0$ である．すなわち，2つの定義に従った$\mathbf{P}_1$に直交する空間は同じであることである．[ヒント：内積の性質**1a. 4**だけを用いればよい．]

19.2　内積の両方の定義に共通にかかわる，$\mathscr{S}$ の正規直交基底が存在する．

19.3　$\mathbf{P}_1, \cdots, \mathbf{P}_m$ は正規直交基底で，$\lambda_i = (\mathbf{P}_i, \mathbf{P}_i)_1/(\mathbf{P}_i, \mathbf{P}_i)_2$ について $\lambda_1 \geqslant \cdots \geqslant \lambda_m$ になる順番に並べられているとする．このとき，λ_i と $\mathbf{P}_i$ を，1つの内積の他の内積に関する固有値と固有ベクトルと呼ぶ．$\mathbf{A}$と$\mathbf{B}$が p. d. 行列なら，$\mathbf{X}'\mathbf{A}\mathbf{Y}$ と $\mathbf{X}'\mathbf{B}\mathbf{Y}$ はともに妥当な内積で，$\lambda_1, \cdots, \lambda_m$ は固有方程式 $|\mathbf{A} - \lambda\mathbf{B}| = 0$ の根となる．

19.4　すべての$\mathbf{X}, \mathbf{Y}$について $\|\mathbf{A}\mathbf{X}\| \leqslant \|\mathbf{A}\mathbf{X} + \mathbf{B}\mathbf{Y}\| \Leftrightarrow \mathbf{A}^*\Lambda\mathbf{B} = 0$，ただし，$(\mathbf{X}, \mathbf{Y}) = \mathbf{Y}^*\Lambda\mathbf{X}$.

19.5　n 次元のユニタリ空間，U^n においては，どの内積も $(\mathbf{X}, \mathbf{Y}) = \mathbf{Y}^*\Lambda\mathbf{X}$ の形に書ける．ただし，Λは p. d. 行列である．

20　p. d. と p. s. d. 行列の性質．

20.1　$\mathbf{A} = (a_{ij})$ が p. d. または p. s. d. で，$a_{ii} = 0$ なら，すべての j について，$a_{ij} = 0$ である．

20.2　$\mathbf{A}$が p. d. で $\mathbf{A}^{-1} = (a^{ij})$ と書くと，

(a)　$a^{ii} \geqslant 1/a_{ii}$，等号は $j = 1, \cdots, i-1, i+1, \cdots, m$ について $a_{ij} = 0$ のときのみ成立する．

(b)　すべての i について $a^{ii} = 1/a_{ii}$ なら，$a_{ij} = 0$，$i \neq j$ となる．

(c)　$|\mathbf{A}| \leqslant a_{11} \cdots a_{pp}$ で，等号は $a_{ij} = 0$，$i \neq j$ のときのみ成立する．

20.3　$\mathbf{A}$を p. d.，$\mathbf{A}_{11}$を正方行列とし，

$$\mathbf{A}^{-1} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix}^{-1} = \begin{pmatrix} \mathbf{A}^{11} & \mathbf{A}^{12} \\ \mathbf{A}^{21} & \mathbf{A}^{22} \end{pmatrix}$$

と書くと，$\mathbf{A}^{11} - \mathbf{A}_{11}^{-1}$ は n. n. d. である．

21　**カントロビッチの不等式**　$\mathbf{A}$は p. d. のエルミート行列で，固有値を $\lambda_1 \geqslant \cdots \geqslant \lambda_m > 0$ とすると，

$$1 \leqslant (\mathbf{X}^*\mathbf{A}\mathbf{X})(\mathbf{X}^*\mathbf{A}^{-1}\mathbf{X}) \leqslant \frac{1}{4}\left[\left(\frac{\lambda_1}{\lambda_m}\right)^{1/2} + \left(\frac{\lambda_m}{\lambda_1}\right)^{1/2}\right]^2$$

ただし，$\mathbf{X}$は規準化してある．すなわち，$\mathbf{X}^*\mathbf{X} = 1$.

22　**1f の結果の拡張**　$\mathbf{A}$を対称行列，$\mathbf{C}$を p. d. 行列とし，$|\mathbf{A} - \lambda\mathbf{C}| = 0$ の根を $\lambda_1 \geqslant \cdots \geqslant \lambda_m$ であらわす．

22.1　$\sup_{\mathbf{X}} \dfrac{\mathbf{X}'\mathbf{A}\mathbf{X}}{\mathbf{X}'\mathbf{C}\mathbf{X}} = \lambda_1$，　$\inf_{\mathbf{X}} \dfrac{\mathbf{X}'\mathbf{A}\mathbf{X}}{\mathbf{X}'\mathbf{C}\mathbf{X}} = \lambda_m$.

22.2　$\inf_{\mathbf{B}} \sup_{\mathbf{B}'\mathbf{C}\mathbf{X}=0} \dfrac{\mathbf{X}'\mathbf{A}\mathbf{X}}{\mathbf{X}'\mathbf{C}\mathbf{X}} = \lambda_{k+1}$，　$\sup_{\mathbf{B}} \inf_{\mathbf{X}\mathbf{B}'\mathbf{C}=0} \dfrac{\mathbf{X}'\mathbf{A}\mathbf{X}}{\mathbf{X}'\mathbf{C}\mathbf{X}} = \lambda_{m-k}$.

ただし，$\mathbf{B}$は $m \times k$ 行列である．

22.3　$\mathbf{X}_1, \cdots, \mathbf{X}_k$ は $\mathbf{X}_i'\mathbf{C}\mathbf{X}_j = 0$，$i \neq j$ をみたすとする．このとき，

$$\sup_{\mathbf{X}_1, \cdots, \mathbf{X}_k} \sum_1^k \frac{\mathbf{X}_i'\mathbf{A}\mathbf{X}_i}{\mathbf{X}_i'\mathbf{C}\mathbf{X}_i} = \lambda_1 + \cdots + \lambda_k,$$

$$\inf_{\mathbf{X}_1, \cdots, \mathbf{X}_k} \prod_1^k \frac{\mathbf{X}_i'\mathbf{A}\mathbf{X}_i}{\mathbf{X}_i'\mathbf{C}\mathbf{X}_i} = \lambda_{m-k+1} \cdots \lambda_m.$$

23　**逆行列とある種の積の計算**　$\mathbf{A}$を次数 m の〔対称〕p. d. 行列，$\mathbf{D}$を $m \times k$ 行列とする．統計的問題では，行列$\mathbf{A}^{-1}$および $\mathbf{D}'\mathbf{A}^{-1}\mathbf{D}$ の計算がしばしば必要である．

23.1　下三角行列$\mathbf{B}$が存在し，$\mathbf{T}$を上三角行列，$\mathbf{A} = \mathbf{T}'\mathbf{T}$ とおくと，$\mathbf{B}\mathbf{A} = \mathbf{T}$ と書けることを示せ．これから，

$$\mathbf{A}^{-1} = \mathbf{B}'\mathbf{B}, \qquad \mathbf{D}'\mathbf{A}^{-1}\mathbf{D} = (\mathbf{B}\mathbf{D})'(\mathbf{B}\mathbf{D})$$

を証明せよ．

23.2　同じ公式は，$\mathbf{A}$が p. s. d. のとき，$\mathbf{A}^-$および $\mathbf{D}'\mathbf{A}^-\mathbf{D}$ の計算にも使える．

23.3　つぎの計算法によって，$\mathbf{T}$と$\mathbf{B}$は同時に得られることを確かめよ．

例示のために，$\mathbf{A}$として次数4の行列（対角線の下の要素を省いてある）をとり，表2に示すように，行列$\mathbf{A}$と単位行列を並べる．

表 2　平方根法による掃き出し

行番号	A				I			
1.1	a_{11}	a_{12}	a_{13}	a_{14}	1			
1.2		a_{22}	a_{23}	a_{24}	·	1		
1.3			a_{33}	a_{34}	·	·	1	
1.4				a_{44}	·	·	·	1
2.1	t_{11}	t_{12}	t_{13}	t_{14}	b_{11}			
2.2		t_{22}	t_{23}	t_{24}	b_{21}	b_{22}		
2.3			t_{33}	t_{34}	b_{31}	b_{32}	b_{33}	
2.4				t_{44}	b_{41}	b_{42}	b_{43}	b_{44}
	T				B			

（i）行 (2.1) は，行 (1.1) の各要素を $t_{11}=\sqrt{a_{11}}$ で割って求める．そのような操作を $(2.1)=(1.1)/t_{11}$ で示す．

（ii）要素 $t_{22}=\sqrt{a_{22}-t_{12}^2}$ とする．行 (2.2) は，$t_{22}\neq 0$ なら $[(1.2)-t_{12}(2.1)]\div t_{22}$ という操作で求める．$t_{22}=0$ なら，$[(1.2)-t_{12}(2.1)]$ として求める．

（iii）要素 $t_{33}=\sqrt{a_{33}-t_{13}^2-t_{23}^2}$ とする．$t_{33}\neq 0$ なら，行 $(2.3)=[(1.3)-t_{13}(2.1)-t_{23}(2.2)]\div t_{33}$ とし，$t_{33}=0$ なら t_{33} の代りに 1 を用いる．同様にして，最後の行 (2.4) を求める．

行列に関連する数値計算については，Dwyer (1951) と Householder (1964) を見よ．

24　$(\mathbf{A}'\mathbf{A})^{-1}$，$\mathbf{D}'(\mathbf{A}'\mathbf{A})^{-1}\mathbf{D}$ などの計算．$\mathbf{A}'$ を階数 k の $m\times k$ 行列，$\mathbf{D}$ を $m\times p$ 行列とする．この $(\mathbf{A}'\mathbf{A})^{-1}$ などを計算する 1 つの方法は，まず $\mathbf{A}'\mathbf{A}$ を計算し，これに問題 **23** の方法を適用することである．しかし，この計算は，$\mathbf{A}'\mathbf{A}$ を計算しないで直接行うことができる．

24.1　[(vi)，**1b.2**] の行列の三角化 $\mathbf{RA}=\begin{pmatrix}\mathbf{T}\\ \mathbf{0}\end{pmatrix}$ を用いよ．ただし，$\mathbf{R}$ は直交行列，$\mathbf{T}$ は上三角行列である．$\mathbf{A}'\mathbf{A}=\mathbf{T}'\mathbf{T}$ で，かつ，$(\mathbf{A}'\mathbf{A})^{-1}=\mathbf{T}^{-1}(\mathbf{T}^{-1})'$ を証明せよ．[三角行列の逆行列 $\mathbf{T}^{-1}$ を計算するのは容易である．また，$\mathbf{D}'(\mathbf{A}'\mathbf{A})^{-1}\mathbf{D}=(\mathbf{D}'\mathbf{T}^{-1})(\mathbf{D}'\mathbf{T}^{-1})'$ である．]

24.2　グラム＝シュミットの標準化 $\mathbf{A}=\mathbf{ST}$ を用いよ．ただし，$\mathbf{S}$ の列は正規直交系である．$\mathbf{A}'\mathbf{A}=\mathbf{T}'\mathbf{T}$ で，かつ，$\mathbf{T}$ は問題 **24.1** の $\mathbf{T}$ と同じであることを示せ．それゆえ，$\mathbf{T}$ を計算する 2 つの方法が得られたが，問題 **24.1** に従って $\mathbf{T}$ を計算する方が，数値的にはより安定であると考えられている．

25　正方行列 $\mathbf{A}$ が行列 $\mathbf{B}$ と相似であるとは，$\mathbf{PAP}^{-1}=\mathbf{B}$ とするような正則行列 $\mathbf{P}$ が存在するときをいう．

（a）相似は $n\times n$ 行列の間の同値関係である．

（b）$\mathbf{A}$ は $\mathbf{B}$ に相似であるとすると，$\mathbf{A}$ と $\mathbf{B}$ は，同じ固有多項式をもち，したがって，固有値もトレースも行列式も同じである．

26　正方行列 $\mathbf{A}$ が半単純であるとは，$\mathbf{P}^{-1}\mathbf{AP}=\Lambda$（対角行列）とするような正則行列 $\mathbf{P}$ が存在するときをいう．Λ の対角要素は $\mathbf{A}$ の固有値で，$\mathbf{P}$ の列はその対応する固有ベクトルである．

$\mathbf{A}$ が半単純であるのは，固有空間が n 次元ユニタリ空間 U^n の全体であるときのみである．

27　$\mathbf{A}$ を $m\times n$ 行列とする．すべての $\mathbf{X},\xi\in U^n$ について，$(\mathbf{X},\xi)=\xi^*\mathbf{NX}$ で，すべての $\eta,\mathbf{Y}\in U^m$ について $(\eta,\mathbf{Y})=\mathbf{Y}^*\mathbf{M}\eta$ とする．ただし，$\mathbf{M}$ と $\mathbf{N}$ は p.d. 行列である．$\mathbf{A}$ の随伴行列 $\mathbf{A}^\#$ を，すべての $\mathbf{X}\in U^n$ と $\mathbf{Y}\in U^m$ についての $(\mathbf{AX},\mathbf{Y})=(\mathbf{X},\mathbf{A}^\#\mathbf{Y})$ という関係で定義する．$m=n$ なら $\mathbf{M}=\mathbf{N}$ とせよ．次を証明せよ．

（i）$(\mathbf{A}^\#)^\#=\mathbf{A}$

（ii）$\mathbf{A}^\#=\mathbf{N}^{-1}\mathbf{A}^*\mathbf{M}$

（iii）$(\mathbf{AB})^\#=\mathbf{B}^\#\mathbf{A}^\#$

（iv）$(\mathbf{A}^\#)^{-1}=(\mathbf{A}^{-1})^\#$

（v）$\mathbf{A}$ が半単純で実数の固有値をもつときのみ，内積を適当に選べば，$\mathbf{A}=\mathbf{A}^\#$ が成り立つ．

28　$\mathbf{A}=\mathbf{U}\begin{pmatrix}\Delta & 0\\ 0 & 0\end{pmatrix}\mathbf{V}^*$ は $\mathbf{A}$ の特異値分解であるとせよ．ただし，$\mathbf{U}$ と $\mathbf{V}$ はユニタリ行列，Δ は正の対角要素をもつ対角行列である．このとき，

（a）$\mathbf{L},\mathbf{M},\mathbf{N}$ をそれぞれ適当な次数の任意の行列として，$\mathbf{G}=\mathbf{V}\begin{pmatrix}\Delta^{-1} & \mathbf{L}\\ \mathbf{M} & \mathbf{N}\end{pmatrix}\mathbf{U}^*$ と書けるときのみ，$\mathbf{G}$ は $\mathbf{A}$ の一般逆行列となる．

（b）$\mathbf{G}=\mathbf{A}_r^-\Leftrightarrow\mathbf{G}=\mathbf{V}\begin{pmatrix}\Delta^{-1} & \mathbf{L}\\ \mathbf{M} & \mathbf{M}\Delta\mathbf{L}\end{pmatrix}\mathbf{U}^*$，　$\mathbf{L}$ と $\mathbf{M}$ は任意．

（c）$\mathbf{G}=\mathbf{A}_m^-\Leftrightarrow\mathbf{G}=\mathbf{V}\begin{pmatrix}\Delta^{-1} & \mathbf{L}\\ 0 & \mathbf{N}\end{pmatrix}\mathbf{U}^*$，　$\mathbf{L}$ と $\mathbf{N}$ は任意．

（d）$\mathbf{G}=\mathbf{A}_l^-\Leftrightarrow\mathbf{G}=\mathbf{V}\begin{pmatrix}\Delta^{-1} & 0\\ \mathbf{M} & \mathbf{N}\end{pmatrix}\mathbf{U}^*$，　$\mathbf{M}$ と $\mathbf{N}$ は任意．

（e）$\mathbf{G}=\mathbf{A}^+\Leftrightarrow\mathbf{G}=\mathbf{V}\begin{pmatrix}\Delta^{-1} & 0\\ 0 & 0\end{pmatrix}\mathbf{U}^*$.

29　$\mathbf{A}$ を正方行列，λ を $\mathbf{A}$ の 1 つの固有値とする．λ に対応する $\mathbf{A}$ のすべての固有ベクトルの集合は，零ベクトルとあわせて部分空間をつくる．この部分空間を，λ に対応する $\mathbf{A}$ の固有空間とよぶ．この部分空間の次元を λ の幾何学的重複度とよぶ．方程式 $|\mathbf{A}-\lambda\mathbf{I}|=0$ の根としての λ の重複度は λ の代数的重複度とよばれる．以下を証明せよ．

（a）代数的重複度 $\geqq$ 幾何学的重複度，各固有値について．

（b）代数的重複度＝幾何学的重複度，$\mathbf{A}$ のすべての固有値について $\Leftrightarrow$ $\mathbf{A}$ は半単純．

（c）$\mathbf{A}$ が実対称行列のときは，各固有値について，その代数的重複度は幾何学的重複度に等しい．

30　$\mathbf{T}=\mathbf{G}+\mathbf{XUX}'$ とし，$\mathbf{G}$ は n.n.d.，$\mathbf{U}$ は対称行列で，$\mathscr{M}(\mathbf{G})\subset\mathscr{M}(\mathbf{T})$ かつ，$\mathscr{M}(\mathbf{X})\subset\mathscr{M}(\mathbf{T})$ とする．このとき，$\mathbf{T}^-$ を任意の一般逆行列とすると，$R(\mathbf{X}'\mathbf{T}^-\mathbf{X})=R(\mathbf{X}')$．[適当な Λ をとると，$\mathbf{X}=(\mathbf{G}+\mathbf{XUX}')\Lambda$，これから，$\mathbf{X}'\mathbf{T}^-\mathbf{X}=\mathbf{X}'\Lambda=\Lambda'\mathbf{X}$ を得ることに注意せよ．このとき，

$$\begin{aligned}\lambda'\mathbf{X}'\mathbf{T}^-\mathbf{X}&=\lambda'\mathbf{X}'\Lambda=\lambda'\Lambda'\mathbf{X}=0\\&=\lambda'\Lambda'(\mathbf{G}+\mathbf{XUX}')\Lambda=0\Rightarrow\lambda'\Lambda'\mathbf{G}\Lambda=0\Rightarrow\lambda'\Lambda'\mathbf{G}=0.\end{aligned}$$

したがって，$\lambda'\Lambda'(\mathbf{G}+\mathbf{XUX}')=\mathbf{0}=\lambda'\mathbf{X}'$ すなわち，$\lambda'\mathbf{X}'\mathbf{T}^{-}\mathbf{X}=\mathbf{0}\Rightarrow\lambda'\mathbf{X}'=\mathbf{0}$ を得たことになる．それゆえ，$\mathscr{M}(\mathbf{X}')\subset\mathscr{M}(\mathbf{X}'\mathbf{T}^{-}\mathbf{X})$．この逆向きも明らかである．]

31　$\mathbf{A}=\begin{pmatrix}\mathbf{A}_{11} & \mathbf{A}_{12}\\ \mathbf{A}_{21} & \mathbf{A}_{22}\end{pmatrix}$ は p. s. d. 行列とすると，$\mathscr{M}(\mathbf{A}_{21})\subset\mathscr{M}(\mathbf{A}_{22})$ で，したがって一般逆行列をどのように選んでも $\mathbf{A}_{22}\mathbf{A}_{22}^{-}\mathbf{A}_{21}=\mathbf{A}_{21}$ が成り立つ．

32　**シュールの補題**　$\mathbf{A}=(a_{ij})$ と $\mathbf{B}=(b_{ij})$ は n. n. d. 行列とする．そのアダマール積 $\mathbf{C}=(a_{ij}b_{ij})$ もまた n. n. d. であることを示せ．

33　$\mathbf{B}$ を階数 k の $k\times p$ 行列，$\mathbf{C}$ を階数 $p-k$ の $(p-k)\times p$ 行列とし，$\mathbf{BC}'=\mathbf{0}$ とする．このとき，

$$\mathbf{C}'(\mathbf{C}\Sigma\mathbf{C}')^{-1}\mathbf{C}+\Sigma^{-1}\mathbf{B}'(\mathbf{B}\Sigma^{-1}\mathbf{B}')^{-1}\mathbf{B}\Sigma^{-1}=\Sigma^{-1}.$$

（この結果は，両辺に $\mathbf{B}$ と $\mathbf{C}\Sigma$ をかけると得られる．）

参　考　文　献

Aitken, A. C. (1944), *Determinants and Matrices* (third edition), University Mathematical Texts, 1, Oliver and Boyd, Edinburgh and London.

Atkinson, F. A., G. A. Watterson, and P. A. P. Moran (1960), A matrix inequality, *Quart. Jour. Math.* (Oxford series) **11**, 137–140.

Anderson, T. W. and S. Dasgupta (1963), Some inequalities on characteristic roots of matrices, *Biometrika* **50**, 522–524.

Berge, Claude (1963), *Topological Spaces* (English first edition), Oliver and Boyd, Edinburgh and London.

Dwyer, P. S. (1951), *Linear Computations*, John Wiley, New York.

Fisher R. A. (1949), *The Theory of Inbreeding*, Oliver and Boyd, Edinburgh and London.

Gantmacher, F. R. (1959), *The Theory of Matrices*, Vols. I and II, Chelsea, New York.

Guttman, Louis (1954), Some necessary conditions for common factor analysis, *Psychometrika* **19**, 149–161.

Halmos, P. R. (1951), *Introduction to Hilbert Space and the Theory of Spectral Multiplicity*, Chelsea, New York.

Householder, A. S. (1964), *The Theory of Matrices in Numerical Analysis*, Blaisdell, New York.

Khatri, C. G. (1968), Some results for the singular multivariate regression models, *Sankhyā A*, 30, 267–280.

Levi, F. W. (1942), *Algebra*, Calcutta University Publication, Calcutta.

Moore, E. H. (1935), *General Analysis*, American Phil. Soc., Philadelphia.

Mulholland, H. P. and C. A. B. Smith (1959), An inequality arising in genetical theory, *Am. Math. Monthly* **66**, 673–683.

Penrose, R. (1955), A generalized inverse for matrices, *Proc. Camb. Phil. Soc.* **51**, 406–413.

Turnbull, H. W. and A. C. Aitken (1932), *An Introduction to the Theory of Canonical Matrices*, Blackie, London.

［著者の本と論文の引用は巻末にのせてある．］

第2章　確率論，道具と手法

まえがき　この章では確率論の基本的概念と統計的推測で使われる道具とについて述べる．集合の代数と測度論の基礎知識があれば，**2a** における確率論の公理的な展開は十分に理解できる．

この章の後半にでてくる測度論のさらに進んだ結果，すなわち，スティルチェス積分，ルベーグの単調収束および上限収束定理，ラドン＝ニコディムの定理などは，付録に述べてある．確率論に応用するさいには，それらの定理を正確に理解していることが必要である．

この章では数理統計を学ぶために必要なほとんどすべての基本的な結果を述べる．ここで述べた大部分の命題についてはくわしい証明をしてある．読者が，数学的な議論をよりよく理解するには，証明のいろいろな段階を学習するとよいであろう．

確率測度族の不変性，統計的手続きの不変性，確率空間における部分集合族，統計量によって定められる部分集合体と，部分集合体に関する条件つき期待値などのすすんだ概念は，この章の終りの付録 **2C** と **2D** とで論ずる．

確率論の極限定理の多くはこの章に入っている．そのうちのいくつかは証明が行われ，それ以外は補足と例題の節に引用されている．また，特性関数もかなりくわしく扱っている．

この章の内容は，さらに高度の確率論および確率過程の特殊な領域を研究しようとする読者にとって適当な準備となるだろう．

さらに研究を進めるためには，読者は，以下の著者たちの本，すなわち，Doob(1953), Feller(1966), Gnedenko and Kolmogorov(1954), Grenander(1963), Kagan, Linnik and Rao(1972), Kolmogorov(1953, 1950), Lévy(1937), Loève(1955), Linnik(1960), Parthasarathy(1967), Ramachandran(1967) などを参照できよう．

2a　確率の計算

2a.1　基本事象の空間

自然界あるいは実験室内での実験において出現する事象を確実に予測することはできない．しかし，ある種の事象は他の種類の事象よりも多く出現し，したがって事象にたいして正確な尺度あるいは指数を与えることによって（つまり確率と呼ばれるものを与えることによって）そのような現象を記述できないかどうかという自然な疑問が生まれる．このような数値が，実用上有用であるためには，直観に基づくいくつかの首尾一貫性の条件をみたさなければならない．確率論はこれらの量をつくりあげ，研究するためのものである．確率論が現実世界の問題に応用できるかどうかということ，すなわち，偶然的な事情のもとでの仮説的な確率の適切さを検討することや特有な確率を推定すること，あるいはその確率に基づいて事象を予測することは，統計理論および統計的方法の扱うべき題目であり，この本の主として扱うところである．

まず最初の段階として，実験のすべての結果のうちある意味で識別可能なものを列挙する．そしてそれらを**基本事象**と呼ぶが，これはある種の実際上の理由から，事象を基本事象に基づく，より一般的な規定として（すなわち，できれば基本事象の集合として）明確に定義したいからである．

いま $1, 2, \cdots, 10$ という番号を記した10個の札があるものとし，これらのうちから1つが無作為にとらえられるものとする．どの番号の札も出現できるのであるから，基本事象は $1, 2, \cdots, 10$ という数字と対応づけられる．とられた数字が素数かどうかを問題にすることができるが，これは基本事象の集合に関する記述であり，基本事象の1つが起こったとき，事象の記述すなわち“とられた番号は素数である”ということを確かめられるのである．われわれは数字の特性すなわち素な番号であるかどうかを考えるときには，二者択一あるいは2つの事象（素であるか素でないか）しか認められないということを証しうる．しかし，そのような記述を行っても，とられた数字を特徴づけるために実行可能で興味のあるいろいろな方法，たとえばその番号が奇数か偶数かを明らかにするような方法が生み出されてくるわけではない．むしろ，事象を基本事象の集合と考えるほうがより一般的である．いま考えている例で，1個の数からなる集合，2個の数からなる集合，… を考えると集合の数は

$$\binom{10}{1}+\binom{10}{2}+\cdots+\binom{10}{10}=2^{10}-1$$

だけある．ここで架空の事象としてとられた札に番号が記されていないということもつけ加え，これを空集合と呼ぶことにすると合計で 2^{10} 事象が可能となる．すべての部分集合に興味があるわけではないが，理論を展開するためには，矛盾のない計算法が構成できる部分集合の全体を考察する必要がある．

次の例として母集団から抽出した個体の身長を考えよう．観測値は理論的に実直線上の連続値を仮定してもよいし，もし身長を 0.5 cm の間隔で測定するならば基本事象は実直線上の可算個の点の集合になる．この場合には興味のある事象は特定の区間に属する身長の個人が現われることであってもよい．このような部分集合の数は有限ではないが，確率が定義できるような部分集合の可能なもっとも広範囲な集合を調べることは意味あることであろう．

2a.2　部分集合（事象）の族

すべての基本事象の集合を Ω とする．これは**標本空間**ともよび，その元または基本事象を要素または点 ω であらわす．もし，事象（部分集合）A が元のもつ特性 π（2a.1 の例にあるように素数というような特性）で定義されるならば，特性 π をもたないような事象を考えることも自然である．この場合，これに対応する部分集合は A に属さない Ω の元のすべての集合である．そのような集合を，A の**補集合**とよび，A^c であらわす．Ω の補集合は空集合で，Ω の元をなにも含まず，$\varnothing$ であらわす．いま，A_1 と A_2 とをそれぞれ特性 π_1, π_2（たとえば，それぞれ素数，偶数というような特性）をもつ2つの事象であるとするとき，2つの他の自然な事象を定義することができる．すなわち，1つは π_1 または π_2 という特性をもつ事象（素数または偶数である番号），もう1つは π_1 かつ π_2 という特性をもつ事象（素数でしかも偶数である番号）である．前者は，A_1 または A_2 に属するすべての元の集合で，$A_1 \cup A_2$ であらわし，**和集合**という．後者は，A_1 と A_2 に属する元の集合で，$A_1 \cap A_2$ であらわし，**共通集合**という．確率が定義される集合の族 C は少なくとも次の性質をもたなければならない．すなわち，$A_1, A_2 \in C$ ならば，A_1^c，$A_1 \cup A_2$，$A_1 \cap A_2 \in C$ という性質であって，これによって有用な理論が生まれるのである．そのような集合族を**ブール集合体**とよび $\mathscr{F}$ であらわす．

これ以外の集合の演算，すなわち，差 $A_1 - A_2$，すなわち，A_1 の元であるが A_2 の元でないものの集合，$A_1 - A_2 = A_1 \cap A_2^c$ を定義することも有用なことである．もちろん，$A_1, A_2 \in \mathscr{F}$ ならば $A_1 - A_2 \in \mathscr{F}$ である．

2a.3　集合関数としての確率

$\mathscr{F}$ に属する集合 A にたいして値 $P(A)$ を与え，A の**確率**とよぶ．いいかえれば，$\mathscr{F}$ の

要素にたいし，集合関数Pを定義するのである．もし，確率の概念を事象の起こる頻度に関係させねばならないとするならば，集合関数Pはある直観的な条件をみたさねばならない．まず第1に，$P(A)$ のとりうる値の幅は $[0,1]$ であって0の値をとることは，事象の起きることが不可能であることを意味し，1は確実に起こることを意味している．第2に，$A_1, A_2, \cdots, A_k$ $(A_i \in \mathscr{F})$ を，それらの和が Ω であるような，互いに素な集合であるとする．これは任意の基本事象が起こったとすると，k個の可能な事象 $A_1, \cdots, A_k$ のうちの1つしかもただ1つの事象が起こったことを意味する．事象 $A_1, \cdots, A_k$ の相対頻度は，合計したときに1になるはずであり，したがって，$P(A_1) + \cdots + P(A_k) = 1$ という条件をみたす．

$A_i \in \mathscr{F}$ $(i = 1, 2, \cdots)$ が，$\bigcup A_i = \Omega$ となるような互いに素な可算個の集合であるとする．$\sum_1^\infty P(A_i)$ はどうなるか？　Ωのどんな有限分解にたいしても $\sum P(A_i) = 1$ になるという第2の直観的な要請からは $\sum_1^\infty P(A_i) \leqslant 1$ が得られるのみである．$\sum_1^\infty P(A_i) < 1$ である場合を調べるために事象 $B_k = \bigcup_k^\infty A_i$ の列を考える．集合 $B_1, B_2, \cdots$ は極限が空集合 $\varnothing$ になるような減少列である．$P(B_k)$ は，kが大きくなると減少して0になると考えてもよい．しかしこれは，$\sum_1^\infty P(A_i) < 1$ なら起こらない．したがって，Ωの有限または可算分解にたいして $\sum P(A_i) = 1$ という条件をつけると便利なのである．以上の直観的な条件は，ある集合族$\mathscr{F}$の上の集合関数Pについての2つの基本公理のかたちで次のように述べることができる．

公理 1.　$P(A) \geqslant 0,\ A \in \mathscr{F}.$

公理 2.　$\bigcup_{i=1}^\infty A_i = \Omega$，任意の $i \neq j$ において $A_i \cap A_j = \varnothing \Rightarrow \sum_1^\infty P(A_i) = 1.$

$\mathscr{F}$のすべての集合にたいして定義され，公理 1, 2 をみたしている集合関数Pを確率測度という．次の結果は公理から導かれるものである．

（i）$0 \leqslant P(A) \leqslant 1,\ A \in \mathscr{F}.$

（ii）$P(\varnothing) = 0$（公理2から $\varnothing \cup (\bigcup_1^\infty A_i) = \Omega = \bigcup_1^\infty A_i$ を考えれば導ける）．

したがって公理2はΩの有限分解にも適用される．

（iii）$P(\Omega) = 1$　$[\Omega \cup \varnothing \cup \varnothing \cdots = \Omega$ であるから$]$

（iv）$\mathscr{F}$の互いに素な任意の可算個の集合について，その和集合もまた$\mathscr{F}$に属するならば，$P(\bigcup A_i) = \sum P(A_i)$ である．

$$(\bigcup A_i) \cup (\bigcup A_i)^c = \Omega = (\bigcup A_i)^c \cup A_1 \cup A_2 \cdots$$

$$P(\bigcup A_i) + P[(\bigcup A_i)^c] = 1 = P[(\bigcup A_i)^c] + P(A_1) + P(A_2) + \cdots.$$

このことから上の結果が得られる．

（v）$A_1 \cap A_2 = \varnothing$ なら $P(A_1 \cup A_2) = P(A_1) + P(A_2)$　（(iv) の特別な場合）．

（vi）A_iを$\mathscr{F}$における非増加列の集合で $\lim_{i\to\infty} A_i = \bigcap A_i \in \mathscr{F}$ とする．このとき

$$\lim_{i\to\infty} P(A_i) = P(\lim_{i\to\infty} A_i).$$

補集合 A_i^cは非減少列で

$$\bigcup A_i^c = (\bigcap A_i)^c = A_1^c \cup (A_2^c - A_1^c) \cup (A_3^c - A_2^c) \cdots$$

である．公理2の (iv) の結果を使うと

$$\begin{aligned} P[(\bigcap A_i)^c] &= P(A_1^c) + P(A_2^c - A_1^c) + \cdots \\ &= P(A_1^c) + [P(A_2^c) - P(A_1^c)] + \cdots \\ &= \lim_{i\to\infty} P(A_i^c) \end{aligned}$$

であることがわかる．これより

$$1 - P(\bigcap A_i) = \lim_{i\to\infty}[1 - P(A_i)]$$

または

$$P(\lim_{i\to\infty} A_i) = P(\bigcap A_i) = \lim_{i\to\infty} P(A_i) \tag{2a. 3. 1}$$

が得られる．同様に，$A_i,\ i = 1, 2, \cdots$ を $\lim A_i = \bigcup A_i \in \mathscr{F}$ となるような$\mathscr{F}$の非減少集合とすると，

$$\lim_{i\to\infty} P(A_i) = P(\bigcup A_i) = P(\lim_{i\to\infty} A_i). \tag{2a. 3. 2}$$

（vii）次も容易に確かめられよう：

（a）$P(A^c) = 1 - P(A)$

（b）$A_1 \subset A_2 \Rightarrow P(A_2 - A_1) = P(A_2) - P(A_1)$

$\Rightarrow P(A_1) \leqslant P(A_2)$

（c）$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2).$

これより不等式 $P(A_1 \cup A_2) \leqslant P(A_1) + P(A_2)$ が得られ，また帰納法によって任意の集合 $A_1, A_2, \cdots$ にたいして $P(\bigcup A_i) \leqslant \sum P(A_i)$ なることが得られる．

（viii）公理 1, $\Leftrightarrow$ (i), (iii), (iv) $\Leftrightarrow$ (i), (iii), (v), (vi).

Kolmogorov (1933) は元来確率関数を規定する公理として (i), (iii), (v), (vi) を述べた．(viii) の等値性は容易に証明できる．

2a. 4　ボレル集合体（σ 集合体）と確率測度の拡張

ブール集合体は，可算個の集合全部の和集合を含んでいる必要はない．そのような和集合のすべて（したがって可算個の共通集合も）を含んでいる集合体をボレル集合体または

σ 集合体とよぶ．所与の集合体 $\mathscr{F}$（または集合の任意の集合）にたいして，$\mathscr{F}$ を含む最小のボレル集合体が存在する．いまそれを $\mathscr{B}(\mathscr{F})$ であらわすと，これは次のようにして示される．すなわち，ボレル集合体は少なくとも1つ存在する．たとえば，Ω のすべての集合の族を考えればよい．これは $\mathscr{F}$ を含む．さらにボレル集合体の任意の共通集合もまたボレル集合体である．このことから $\mathscr{F}$ を含むすべてのボレル集合体の共通集合はまさしく最小のボレル集合体 $\mathscr{B}(\mathscr{F})$ である．すると，$\mathscr{F}$ 上で定義され，公理 1, 2 をみたす集合関数は $\mathscr{B}(\mathscr{F})$ のすべての集合にたいして一意的に拡張することができる．すなわち，次のような一意関数 P^* が存在する．

（a）$0 \leqslant P^*(A) \leqslant 1$，$A \in \mathscr{B}(\mathscr{F})$.

（b）$\mathscr{B}(\mathscr{F})$ のなかの互いに素な可算個の集合 A_n にたいして $P^*(\bigcup A_n) = \sum P^*(A_n)$.

（c）$A \in \mathscr{F}$ ならば，$P^*(A) = P(A)$.

関数 P^* は次のように定義される．いま $\mathscr{B}(\mathscr{F})$ 内の集合 A と $A \subset \bigcup_1^\infty A_i$ となるような $\mathscr{F}$ 内の集合 A_i の集合の全体を考える．このとき

$$P^*(A) = \inf_{A_i} \sum P(A_i).$$

P の拡張 P^* は測度論の基礎的な結果である（Halmos (1950), Munroe (1959)）．

このより広い集合体 $\mathscr{B}(\mathscr{F})$ を確率関数が定義される基本集合族と考えることは，$\mathscr{B}(\mathscr{F})$ を生成するブール集合体上に関数を定義すれば十分であることから，有用であるといえよう．このようにして元 ω（基本事象）の基本空間 Ω および Ω における集合（事象）のボレル集合体または σ 集合体 $\mathscr{B}$，さらに $\mathscr{B}$ 上の確率測度 P を基礎にして確率論をつくりあげることができることになる．これらの3個を組にした $(\Omega, \mathscr{B}, P)$ を**確率空間**とよび，Ω または $(\Omega, \mathscr{B})$ を**標本空間**とよぶ．

2a.5　確率変数と分布関数の概念

ここで，この章を通じて使ういくつかの記号を紹介しておくことにしよう．すでに確率空間 $(\Omega, \mathscr{B}, P)$ は定義した．実直線 $(-\infty, \infty)$ は R であらわし，また，n 次元実ユークリッド空間は R^n であらわす．R^n の点は順序のつけられた実数座標値の集合 $(x_1, x_2, \cdots, x_n)$ であらわせる．これらは Ω が特別な空間の場合である．

R^n において，開集合，閉集合，あるいは n 次元の閉直方体集合，開直方体集合，あるいは

$$a_i \leqslant x_i < b_i, \qquad \text{ただし，} -\infty \leqslant a_i \leqslant b_i \leqslant \infty, \qquad i = 1, \cdots, n$$

のような形で与えられる集合の有限和集合によって生成される集合体の族 $\mathscr{G}$ を考える．$\mathscr{G}$ を含む最小ボレル集合体はおのおのの場合同じで $\mathscr{B}_n$ であらわす．確率空間 $(R^n, \mathscr{B}_n, P)$ は実際上の応用で重要な役割を演ずる．R^∞ と $(R^\infty, \mathscr{B}_\infty, P)$ に関係する $\mathscr{B}_\infty$ の概念は 2c.1 で説明する．

確率変数　空間 $(\Omega, \mathscr{B}, P)$ 上で定義された実数値点関数 $X(\cdot)$ が，もし R に属するすべての x にたいして $\{\omega : X(\omega) < x\} \in \mathscr{B}$ ならば，これを**確率変数** r.v. と呼ぶ（測度論のことばで可測という）．

関数 X は Ω から R のなかへの写像とみなすことができ，$\Omega \xrightarrow{X} R$ と書ける．したがって，R の点の集合 S にたいして，Ω のなかの点の集合で S へ写像されるものが存在し，それは S の原像と呼び $X^{-1}(S)$ であらわす．この記法を使えば，X は

$$\text{すべての } x \text{ にたいして} \quad X^{-1}(-\infty, x) \in \mathscr{B}$$

であるときそしてそのときのみ確率変数である．

分布関数　R の上で定義される点関数

$$F(x) = P\{\omega : X(\omega) < x\} = P[X^{-1}(-\infty, x)]$$

を**分布関数** d.f. と呼ぶ．このように定義された分布関数はいくつかの重要な性質をもっている．以下 $P(X < x)$ を $P\{\omega : X(\omega) < x\}$ という意味で使う．

（i）$\underline{F(-\infty) = 0,\ F(\infty) = 1.}$

$x \to -\infty$ のとき $X^{-1}(-\infty, x)$ の極限は $\varnothing$，すなわち $\mathscr{B}$ における空集合である．このことから $x \to -\infty$ のとき $\lim F(x) = 0$，つまり $F(-\infty) = 0$ である．これは，そうでないと (2a.3.1) に矛盾するので明らかである．同様に $F(\infty) = 1$ である．

（ii）$\underline{F \text{は非減少である．}}$

$x_1 < x_2$ とする．$(-\infty, x_2) \supset (-\infty, x_1)$ であることから $X^{-1}(-\infty, x_2) \supset X^{-1}(-\infty, x_1)$ で，したがって $P[X^{-1}(-\infty, x_2)] \geqslant P[X^{-1}(-\infty, x_1)]$ である．

（iii）$\underline{F \text{は少なくとも左側から連続である．}}$

いま，$x_1, x_2, \cdots$ を増加数列，$A_1, A_2, \cdots$ を $\mathscr{B}$ でこれらに対応する増加集合列とする．もし $\lim x_n = x$ ならば $\lim A_n = \bigcup A_n = A$ である．ただし $A = X^{-1}[(-\infty, x)]$ である．(2a.3.2) を使って次の式を得る．

$$\lim F(x_n) = \lim P(A_n) = P(\lim A_n) = P(A) = F(x).$$

ここで，上からの極限を考えれば x の原像はすべての A_i に属し，かつ，その極限集合は $X^{-1}((-\infty, x])$ で，その確率は $F(x) + P(X = x)$ であることに注意せよ．したがって，不連続性はゼロでない確率をもつ1点の集合において発生しうる．

左側連続性は $F(x) = P(X < x)$ という定義の結果である．著者によっては分布関数 $F(x)$ を $P(X \leqslant x)$ と定義する．このときには右連続となる．いずれの定義に従っても その値は，すべての連続点では一致し，不連続点では $P(X = x)$ だけ異なる．2つの定義が同等であり，そのいずれを選ぶかは便宜上の問題であることがわかったから，ここで

は $F(x)=P(X<x)$ を採用する．

（iv）Fの不連続点の集合Sはたかだか可算個である．

Fが $1/n$ より小さくない（$\geqslant 1/n$）値だけ大きくなる点の集合を S_n によってあらわす．このとき $S=\bigcup S_n$．すべてのnにたいして集合 S_n は有限である．なぜなら，そうでないとFは1を超える．このことからSはたかだか可算個である．

（v）もし，2つの分布関数がRにおける稠密な集合の上で一致するならば，それは同一である．

（vi）分布関数の分解．$F(x)=F_1(x)+F_2(x)$ ただし F_2 はいたるところで連続で，F_1 は階段関数とする．このような分解は一意的である．

いま $D=(x_1, x_2, \cdots)$ を $F(x)$ の不連続点とし，

$$F_1(x)=\sum_{x_i<x}[F(x_i+0)-F(x_i)]$$

と定義する．明らかに $F_1(x)$ は階段関数である．このとき，$F_2(x)=F(x)-F_1(x)$ は連続関数であることが容易にわかる．連続部分を

$$F_2(x)=F_{21}(x)+F_{22}(x)$$

とあらわすような，より一般的な分解が存在する．ここで F_{21} は絶対連続すなわちある関数の不定積分であらわされるものであり，F_{22} はほとんどいたるところで導関数がゼロになるような関数である．証明は Loève (1955) を見よ．

（vii）$(\Omega, \mathscr{B}, P)$ 上の確率変数 $X(\cdot)$ から $(R, \mathscr{B}_1)$ 上の確率測度が導かれる．確率測度は $F(x)$ によって完全に指定される．

$S\in\mathscr{B}_1$ とする．$X^{-1}(S)\in\mathscr{B}$ は容易にわかる．$Q(S)=P[X^{-1}(S)]$ と定義する．明らかに $Q(S)$ は，基本公理1,2から証明されるように $\mathscr{B}_1$ 上の確率測度であって

$$S=(-\infty, x) \text{ とすると，} Q(S)=F(x) \qquad (2a.5.1)$$

という性質をもつ．つぎに，$Q'(S)$ を性質 (2a.5.1) をもつ $\mathscr{B}_1$ 上の別の確率測度とすると，すべての $S\in\mathscr{B}_1$ にたいして $Q(S)=Q'(S)$ となることを示そう．そうすれば，(vii) は証明されたことになる．

いま $S=[a, b)$ とすると，$(-\infty, b)=(-\infty, a)+[a, b)$ であることから

$$F(b)=F(a)+Q([a, b))=F(a)+Q'([a, b)).$$

このことから Q と Q' は $[a, b)$ という形の区間で一致する．

$(-n, n)$ 内の $[a, b)$ の形をした集合の有限和集合の族を $\mathscr{F}_n$ とする．このとき $\mathscr{F}_n$ は集合体で，集合 $S\in\mathscr{F}_n$ は $S_i=[a_i, b_i)$ の形の互いに素な集合の和集合でいつもあらわされる．したがって

$$Q(S)=\sum Q([a_i, b_i))=\sum Q'([a_i, b_i))=Q'(S)$$

であり Q と Q' は $\mathscr{F}_n$ 上で，したがってまた $\mathscr{B}(\mathscr{F}_n)$ 上で一致する．

さて，$S\in\mathscr{B}_1$ とすると，$S\cap[-n, n)\in\mathscr{B}(\mathscr{F}_n)$．したがって $Q[S\cap[-n, n)]=Q'[S\cap[-n, n)]$ であり，$n\to\infty$ とすれば $Q(S)=Q'(S)$．

（viii）確率変数 $X(\cdot)$ は $\mathscr{B}_X\subset\mathscr{B}$ なる σ 集合体 $\mathscr{B}_X$ を定義する．すなわち，$\mathscr{B}_X$ は ω 集合の部分 σ 集合体である．

$S\in\mathscr{B}_1$ とし $\mathscr{B}_X=\{X^{-1}(S): S\in\mathscr{B}_1\}$ を考える．$\mathscr{B}_X$ は σ 集合体である．なぜなら

（a）$X^{-1}(-\infty, \infty)=\Omega$，すなわち $\Omega\in\mathscr{B}_X$．

（b）$X^{-1}(S^c)=[X^{-1}(S)]^c$，すなわち $\mathscr{B}_X$ 内の集合の補集合は $\mathscr{B}_X$ に属する．

（c）$X^{-1}(\bigcup_1^\infty S_i)=\bigcup_1^\infty X^{-1}(S_i)$，すなわち $\mathscr{B}_X$ 内の集合の可算和集合は，また $\mathscr{B}_X$ に属する．$\mathscr{B}_X\subset\mathscr{B}$ は明らかである．

（ix）確率変数のボレル関数はまた確率変数である．

実直線Rの点からRのなかへの写像である関数fが次の条件をみたすときボレル関数という：

$$S\in\mathscr{B}_1 \Rightarrow f^{-1}(S)\in\mathscr{B}_1.$$

さらに一般に，n 次元空間 R^n の点からRのなかへの写像である関数fが

$$S\in\mathscr{B}_1 \Rightarrow f^{-1}(S)\in\mathscr{B}_n$$

ならばボレル関数という．

いま X を確率変数とすれば $X^{-1}(S)\in\mathscr{B}$ である．$f(X(\omega))=g(\omega)$ を考えると $g^{-1}(S)=X^{-1}[f^{-1}(S)]\in\mathscr{B}$ であることがわかる．

（x）次の性質をもつ任意の $F(x)$ について：

（a）$F(x)$ は非減少である

（b）$F(-\infty)=0$，$F(\infty)=1$

（c）$F(x)$ は少なくとも左から連続．

そのようなFにたいして，これを分布関数とする確率変数が定義される．

$\Omega=[0, 1]$ とし，$[0, 1]$ の区間によって生成されるボレル集合を $\mathscr{B}$ とし，$P([0, \omega))=\omega$，$0\leqslant\omega\leqslant 1$ であるようなルベーグ測度をPとしよう．このとき確率変数

$$X(\omega)=\inf_{F(y)>\omega} y, \qquad (0\leqslant\omega\leqslant 1)$$

は次の分布関数をもつ：

$$P[\omega: X(\omega)<x]=P[\omega: \inf_{F(y)>\omega} y<x]$$
$$=P[\omega: 0\leqslant\omega<F(x)]=F(x).$$

以上のことから適当な確率空間 $(\Omega, \mathscr{B}, P)$ 上で定義された関数（確率変数）から $F(x)$

が生み出されることが示された．

実際には2つの型の分布関数に関心が払われる．第1は階段関数型で，有限個あるいは可算個の離散値を $a_1 < a_2 < \cdots$ とするときそれに対応する確率 $p_1, p_2, \cdots$ $(\sum p_i = 1)$ をもつ確率変数に対応する．分布関数は $a_1, a_2, \cdots$ の点で飛躍があり，区間 $(a_i, a_{i+1}]$ では定数にとどまる．a_i での変化量は p_i である．第2は絶対連続型といわれ，次の式であらわされる：

$$F(x) = \int_{-\infty}^{x} f(x)dx. \tag{2a.5.2}$$

ここで $f(x)$ は確率密度である．$f(x)$ は x で連続なら $F'(x) = f(x)$ である．これら2つの型以外の分布関数もありうるが，これらは上の2つ型の混合型であり (2a.5.2) の式ではあらわせない．

有限個あるいは可算個の異なる状態をもつ離散分布で統計学で有用ないくつかの例を次にあげる．それぞれの例において，確率変数は X，確率 $P(X=r)$ は p_r であらわす．p_r を指定するさいに必要な他の値すなわちパラメータはギリシャ文字であらわす．変数およびパラメータの範囲はそれぞれ示してある．

2項分布

$$p_r = \binom{n}{r}\pi^r(1-\pi)^{n-r},\ 0<\pi<1,\ n\text{は任意の整数},\ r=0,1,\cdots,n \tag{2a.5.3}$$

幾何分布

$$p_r = (1-\pi)\pi^r,\ 0<\pi<1,\ r=0,1,\cdots,\infty \tag{2a.5.4}$$

ポアソン分布

$$p_r = e^{-\lambda}\cdot\lambda^r/r!,\ \lambda>0,\ r=0,1,\cdots,\infty \tag{2a.5.5}$$

負の2項分布

$$p_r = \binom{k+r-1}{r}\frac{\rho^r}{(1+\rho)^{k+r}},\ \rho>0,\ k\geqslant 1,\ r=0,1,\cdots,\infty \tag{2a.5.6}$$

対数分布

$$p_r = \frac{1}{-\log(1-\alpha)}\frac{\alpha^r}{r},\ 0<\alpha<1,\ r=1,2,\cdots,\infty \tag{2a.5.7}$$

超幾何分布

$$p_r = \binom{A}{r}\binom{N-A}{n-r}\div\binom{N}{n},\ A\leqslant N\ (\text{ともに整数}),\ n\leqslant N\ (n\text{は整数}) \tag{2a.5.8}$$

$r = \max(0, n-N+A), \cdots, \min(A, n)$.

連続分布の重要な例は次の密度関数をもつ正規分布である：

$$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma}e^{-(x-\mu)^2/2\sigma^2},\ -\infty<\mu<\infty,\ 0<\sigma<\infty,\ -\infty<x<\infty. \tag{2a.5.9}$$

次のことは簡単に証明できる：

$$\int_{-\infty}^{\infty} f(x)dx = \frac{1}{\sqrt{2\pi}\,\sigma}\int_{-\infty}^{\infty} e^{-(x-\mu)^2/2\sigma^2}dx = \frac{2}{\sqrt{2\pi}\,\sigma}\int_0^{\infty} e^{-y^2/2\sigma^2}dy,$$

$$(\text{ここで } y = x-\mu)$$

$$= \frac{1}{\sqrt{\pi}}\int_0^{\infty} z^{-1/2}e^{-z}dz = \frac{1}{\sqrt{\pi}}\Gamma\left(\frac{1}{2}\right) = 1.$$

その他の興味ある連続分布は第3章で示す．

2a.6　多次元確率変数

Ω から R^m のなかへの写像である m 個の1変数実関数 $X_1(\cdot), \cdots, X_m(\cdot)$ の集まりが，すべての $\mathbf{x}$ にたいして

$$S(\mathbf{x}) = \{\omega : X_1(\omega) < x_1, \cdots, X_m(\omega) < x_m\} \in \mathscr{B}$$

であるとき m 次元確率変数という．ここで $\mathbf{x}$ は $(x_1, \cdots, x_m)$ をあらわす．1次元確率変数の場合と同様に分布関数を

$$P[S(\mathbf{x})] = F(x_1, \cdots, x_m)$$

で定義する．2a.5 の論議と同様に，F は非減少で各変数について少なくとも左から連続であり，$\mathscr{B}_m$ の上の確率測度を定義することが示される．

X_i だけの分布関数は

$$F_i(x) = F(\infty, \cdots, \infty, x, \infty, \cdots, \infty)$$

であらわされるが，右辺で x は i 番目の位置とする．この F_i を X_i の周辺分布という．同様に，変数の任意の部分集合の周辺分布は $F(x_1, \cdots, x_m)$ でほかの変数を ∞ におきかえることによって得られる．

考えられる2つの重要な型は階段関数と絶対連続関数である．階段関数では，$x_1, \cdots, x_m$ のとる値のいろいろな組み合せにたいする確率を各々指定する．絶対連続関数については，密度関数 $f(x_1, \cdots, x_m)$ は

$$F(x_1, \cdots, x_m) = \int_{-\infty}^{x_1}\cdots\int_{-\infty}^{x_m} f(x_1, \cdots, x_m)dx_1\cdots dx_m$$

となるように与えられる．もし $f(x_1, \cdots, x_m)$ がすべての変数に関して連続ならば

$$\frac{\partial^m F}{\partial x_1\cdots\partial x_m} = f(x_1, \cdots, x_m)$$

である．

2a.7　条件つき確率と統計的独立性

2a.1 の10枚の札の例で，r 個の数から成る任意の部分集合に $r/10$ という確率を与えると，それは確率測度の公理1,2をみたす．定義された測度より，事象 A すなわち札の番号

が素数である確率は 4/10，事象Cすなわち札の番号が偶数である確率は 5/10，そして事象 AC（偶数かつ素数）の確率は 1/10 である．

1枚の札をぬきとりその番号が素数であることがわかった，つまり，事象Aが起きたとしよう．このとき次の問題，すなわち，札の番号が偶数（でかつ素数）である確率，つまり事象 AC が起こる確率はいくらか，が問われるとしよう．当然，AC についての不確実さの度合は，Aが起こったという知識の分だけ減って，Aが真である場合の可能性だけを考えればよいことになる．この例では，すべて素数である 2, 3, 5, 7 という結果だけを考えればよく，その部分集合に確率を定めねばならないのである．そのための自然なやりかたは，基本事象 2, 3, 5, 7 にたいしてもとの確率に比例した確率を与えるのである．それは，われわれの知識によって相対頻度は変化しないからである．よって，新しい確率は $P^*(i) \propto P(i)$ とする，つまり，

$$P^*(i) = \frac{P(i)}{P(A)}, \qquad i = 2, 3, 5, 7$$

$$\sum P^*(i) = 1$$

とおく．ここで $P(A) = P(2) + P(3) + P(5) + P(7)$ は番号が素数である確率である．与えられた番号が素数であるとわかっている上でさらに偶数（事象C）である確率は

$$P(C|A) = P^*(2) = \frac{P(2)}{P(A)} = \frac{P(AC)}{P(A)}$$

とおく．$P(A) \neq 0$ であるとき $P(AC)/P(A)$ は A を与えたときの C の条件つき確率といい，$P(C|A)$ であらわす．$P(A) = 0$ のとき $P(C|A) = 0$ とする．

一般に $(\Omega, \mathscr{B}, P)$ 上で，$P(A) \neq 0$ である集合 A と $\mathscr{B}$ に属する AC, C の集合を考え，次式を定義する：

$$P(C|A) = \frac{P(AC)}{P(A)}. \tag{2a. 7. 1}$$

Aを固定し，$\mathscr{B}$ に属するすべての集合 C にたいして定義した集合関数 $P(\cdot|A)$ は公理 1, 2 を満足し，A を与えたときの条件つき確率測度とよぶ．

事象の独立性　もし $P(C|A) = P(C)$ であるならば，事象Cは事象Aと統計的に独立であるとよばれ，その場合

$$P(AC) = P(A)P(C) \tag{2a. 7. 2}$$

が成り立つ．そのとき，$P(A|C) = P(A)$ すなわち，A は C と独立である．$P(A) = 0$ または $P(C) = 0$ である場合も含めて，もしも (2a. 7. 2) が成り立つならば 2 つの事象 A とCとは独立であるという．

事象 $A_1, A_2, \cdots$ が互いに独立であるとは，そのなかから，n 個の事象 $A_{i_1}, \cdots, A_{i_n}$ をどのように選んでも

$$P(A_{i_1} \cdots A_{i_n}) = P(A_{i_1}) \cdots P(A_{i_n}), \qquad n = 2, 3, \cdots$$

が成り立つときをいう．

2a. 8　確率変数の条件つき分布

確率変数の対 (X_1, X_2) と $P(X_1 \in S) \neq 0$ であるような集合 $S \in \mathscr{B}_1$ を考える．条件つき確率の定義 (2a. 7. 1) によって

$$\frac{P[\omega : X_1(\omega) \in S,\ X_2(\omega) < x_2]}{P[X_1(\omega) \in S]} = F(x_2|X_1 \in S) \tag{2a. 8. 1}$$

は $X_1 \in S$ のもとでの $X_2 < x_2$ の条件つき確率であると解釈することができる．(2a. 8. 1) 式を x_2 の関数と考えるとき，それを $X_1 \in S$ のもとでの X_2 の条件つき分布関数とよび，次のようにあらわす：

$$F_2(\cdot\,|X_1 \in S). \tag{2a. 8. 2}$$

しかし，もし F_1 が X_1 の分布関数で，すべての S と x_2 にたいして

$$P(X_1 \in S, X_2 < x_2) = \int_S F_2(x_2|x_1) dF_1(x_1) \tag{2a. 8. 3}$$

が成り立つような関数 $F_2(\cdot|\cdot)$ が存在するなら，$F_2(\cdot|x_1)$ を $X_1 = x_1$ のときの X_2 の条件つき分布関数とよぶことができるであろう．一般の場合にそのような関数の存在を証明するのはむずかしい．（その存在の証明をしている **2b. 3** を参照．そこでは 2 つの確率変数 X_1 と X_2 との標本空間として R^2 を考え，R^n で使われた論法を用いる．）

この本で考える応用の大部分では，しかしながら，絶対連続な分布関数かまたは離散分布に対応する分布しか考えない．そのような場合には条件つき分布を定義することは簡単である．いま $f(x_1, x_2)$ を X_1 と X_2 の同時密度とし，X_1 だけの密度は $f_1(x_1)$ であるとする（離散分布については“密度”は“確率”でおきかえ，積分は総和でおきかえられる）．定義より

$$P(X_1 \in S, X_2 < x_2) = \int_S \int_{-\infty}^{x_2} f(x_1, x_2) dx_1 dx_2$$

$$= \int_S f_1(x_1) dx_1 \int_{-\infty}^{x_2} \frac{f(x_1, x_2)}{f_1(x_1)} dx_2, \qquad f_1(x_1) \neq 0$$

であり，このことから

$$F_2(x_2|x_1) = \int_{-\infty}^{x_2} \frac{f(x_1, x_2)}{f_1(x_1)} dx_2 \tag{2a. 8. 4}$$

は条件つき分布関数の定義をみたす．さらに $f(x_1, x_2)/f_1(x_1)$ は $X_1 = x_1$ のときの X_2 の条件つき確率密度と考えられる．$f_1(x_1) = 0$ のとき条件つき分布関数と確率密度とをどのように定義するかということは重要ではない．

2つの確率変数 X_1 と X_2 は，すべての x_2 と $S \in \mathscr{B}_1$ にたいして

$$F_2(x_2|X_1 \in S) = F_2(x_2) \tag{2a.8.5}$$

が成り立つとき，独立であるという．Sを1つの区間にとると，(2a.8.5) により，

$$F(x_1, x_2) = F_1(x_1)F_2(x_2) \tag{2a.8.6}$$

となり，また，(2a.8.6) から (2a.8.5) が得られることは自明である．密度関数が存在するときは，独立であるための条件は $f(x_1, x_2) = f_1(x_1)f_2(x_2)$ と書ける．

2b　数学的期待値と確率変数の積率

2b.1　数学的期待値の性質

確率変数Xの数学的期待値は平均ともよばれ

$$E(X) = \int x dF \tag{2b.1.1}$$

で定義される．ここで $F(x)$ は X の分布関数であり，積分はルベーグ＝スティルチェス積分で，この積分についてはこの章の付録 **2A** で説明する．X が確率分布関数 $f(x)$ をもつとき，

$$E(X) = \int x dF = \int x f(x) dx \tag{2b.1.2}$$

であってリーマン積分となることがわかる．$F(x)$ が階段関数である離散確率分布にたいして期待値は

$$E(X) = \int x dF = \sum x_r p_r \tag{2b.1.3}$$

のように総和で示される．(2b.1.3) では，各 x_r はFの不連続点で，p_r は x_r における飛躍の値である．上に示した3つのすべての場合に同じ記号 $\int x dF$ が使えると注意できるのは面白い．

$Y = g(X)$ をXのボレル関数とし，Y の分布関数を $G(y)$ であらわすとする．積分の定義を使って，

$$E(Y) = E[g(X)] = \int y dG(y) = \int g(x) dF(x) \tag{2b.1.4}$$

が成り立つことを証明するのは容易である．したがって，確率変数の関数の期待値は，確率変数の関数の分布関数をまずきめなくても直接に計算できる．

さらに一般的に，k 次元確率変数 $(X_1, \cdots, X_k)$ を考え，$Y = h(X_1, \cdots, X_k)$ がボレル関数であるとするならば

$$E(Y) = \int y dG(y) = \int h(x_1, \cdots, x_k) dF(x_1, \cdots, x_k)$$

が示される．次の結果も証明はやさしい．

（i）cは定数として，$E(cX) = cE(X)$.

（ii）X_1, X_2 が互いに独立であるかどうかに関係なく $E(X_1 + X_2) = E(X_1) + E(X_2)$, より一般的に

$$E(c_1X_1 + \cdots + c_kX_k) = c_1E(X_1) + \cdots + c_kE(X_k)$$

である．

（iii）X_1 と X_2 とが互いに独立ならば $E(X_1X_2) = E(X_1)E(X_2)$.

2b.2　積　　率

確率変数Xの，与えられた定数aのまわりのs次の積率（モーメント）およびs次の絶対積率は，次のように定義される：

$$\alpha_s = E(X-a)^s = \int (x-a)^s dF,$$

一般に

$$= \int (x-a)^s f(x) dx, \qquad \text{密度関数が存在するとき}$$

$$= \sum_r (x_r - a)^s p_r, \qquad \text{離散分布について}$$

$$\gamma_s = E|X-a|^s = \int |x-a|^s dF.$$

ここでは γ_s が存在するときのみ α_s が存在するということにする．これは，α_s を定義する積分がルベーグ＝スティルチェス積分（付録 **2A**）であるときには正しい．$a = E(X)$ のときを中心積率といい，μ_s であらわす．中心積率と任意の原点まわりの積率との，次に示す関係は容易に証明できる．

$$\mu_0 = 1, \quad \mu_1 = 0, \quad \mu_2 = \alpha_2 - \alpha_1^2, \quad \mu_3 = \alpha_3 - 3\alpha_1\alpha_2 + 2\alpha_1^3,$$
$$\mu_4 = \alpha_4 - 4\alpha_1\alpha_3 + 6\alpha_1^2\alpha_2 - 3\alpha_1^4$$

であり，一般に

$$\mu_n = \alpha_n - \binom{n}{1}\alpha_1\alpha_{n-1} + \binom{n}{2}\alpha_1^2\alpha_{n-2} - \cdots (-1)^{n-1}(n-1)\alpha_1^n$$

である．これは次のように証明される．

$$\begin{aligned}
\mu_n &= E[X - E(X)]^n = E[X - a - E(X-a)]^n \\
&= E(X - a - \alpha_1)^n \\
&= E\left\{(X-a)^n - \binom{n}{1}(X-a)^{n-1}\alpha_1 + \binom{n}{2}(X-a)^{n-2}\alpha_1^2 - \cdots\right\} \\
&= E(X-a)^n - \binom{n}{1}\alpha_1 E(X-a)^{n-1} + \binom{n}{2}\alpha_1^2 E(X-a)^{n-2} - \cdots
\end{aligned}$$

$$= \alpha_n - \binom{n}{1}\alpha_1\alpha_{n-1} + \binom{n}{2}\alpha_1^2\alpha_{n-2} - \cdots.$$

平均 μ と，一般に σ^2 であらわされる2次中心積率 μ_2（分散）は，統計データの解析で重要な役割を果たす．分布を記述する測度としての観点だけからは，平均は確率変数の中心値をあらわし，分散は中心値のまわりのちらばりをあらわす．分布の平均と分散についての次のようないくつかの結果が得られる．

（i）2次積率 $\alpha_2(a)$ は a が平均 μ のとき最小である．

$$\begin{aligned} E(X-a)^2 &= E(X-\mu+\mu-a)^2 \\ &= E(X-\mu)^2 + (\mu-a)^2 + 2(\mu-a)E(X-\mu) \\ &= E(X-\mu)^2 + (\mu-a)^2 \geqslant E(X-\mu)^2. \end{aligned}$$

（ii）もし X が非負の確率変数であって，$E(X)$ が存在するならば，

$$E(X) = \int [1-F(x)]dx. \tag{2b. 2. 1}$$

が成立する．部分積分を用いて

$$\begin{aligned} \int_0^T x dF(x) &= TF(T) - \int_0^T F(x)dx \\ &= -T[1-F(T)] + \int_0^T [1-F(x)]dx. \end{aligned} \tag{2b. 2. 2}$$

である．一方

$$T[1-F(T)] = T\int_T^\infty dF(x) < \int_T^\infty x dF(x) \to 0, \qquad T \to \infty \text{ のとき}$$

である．したがって，結果 (2b. 2. 1) は (2b. 2. 2) の両辺の極限をとることによって得られる．

X が任意の確率変数で $E(X)$ が存在するならば，

$$E(X) = \int_0^\infty [1-F(x)-F(-x)]dx \tag{2b. 2. 3}$$

が成立し

$$\lim_{x\to-\infty} xF(x) = \lim_{x\to\infty} x[1-F(x)] = 0$$

は $E(x)$ が存在するための必要条件である．

さらに，もし G と F が有限な平均値 g, f をもつ分布関数であるならば

$$\int_{-\infty}^{+\infty} [G(x)-F(x)]dx = f-g \tag{2b. 2. 4}$$

が成り立つ．Savage (1970) は，より一般的な結果として

$$\int_{-\infty}^{+\infty} [(G*H)(x)-(F*H)(x)]dx = f-g \tag{2b. 2. 5}$$

を示した．ここで H は任意の分布関数で，$G*H$ は分布関数のたたみこみをあらわす

((2b. 4. 12) で示す)．

（iii）チェビシェフの不等式　平均 μ と分散 σ^2 をもつ任意の確率変数 X にたいして

$$P(|X-\mu| \geqslant \lambda\sigma) \leqslant \frac{1}{\lambda^2}, \qquad \lambda > 0 \tag{2b. 2. 6}$$

が成り立つ．

これは，X の分布そのものの性質と独立であるから，重要な不等式である．定義から

$$\begin{aligned} \sigma^2 &= \int (x-\mu)^2 dF(x) \\ &= \int_{-\infty}^{\mu-\lambda\sigma} (x-\mu)^2 dF + \int_{\mu-\lambda\sigma}^{\mu+\lambda\sigma} (x-\mu)^2 dF + \int_{\mu+\lambda\sigma}^{\infty} (x-\mu)^2 dF \end{aligned}$$

である．ここで，中央の項を除き第1項と第3項の $(x-\mu)^2$ をその最小値でおきかえると

$$\begin{aligned} \sigma^2 &\geqslant \lambda^2\sigma^2 \int_{-\infty}^{\mu-\lambda\sigma} dF + \lambda^2\sigma^2 \int_{\mu+\lambda\sigma}^{\infty} dF \\ &\geqslant \lambda^2\sigma^2 P(|x-\mu| \geqslant \lambda\sigma) \end{aligned}$$

を得る．これから不等式 (2b. 2. 6) が出る．(2b. 2. 6) は σ^2 を $E(X-\mu)^2$ でおきかえるならば，μ が平均値でないときにも成り立つ．

（iv）X を $E(X)=\mu$ となる非負確率変数とすると

$$P(X \geqslant t\mu) \leqslant \frac{1}{t}.$$

（v）X_1, X_2 をそれぞれ平均 μ_1, μ_2，分散 σ_1^2, σ_2^2，分布関数 F_1, F_2 をもつ2つの確率変数とし，各 x にたいして

$$F_1(x+\mu_1) - F_1(-x+\mu_1) \geqslant F_2(x+\mu_2) - F_2(-x+\mu_2) \tag{2b. 2. 7}$$

ならば $\sigma_1^2 \leqslant \sigma_2^2$ である．

$|X_1-\mu_1|$ と $|X_2-\mu_2|$ の分布関数をそれぞれ G_1, G_2 とする．部分積分を用いて

$$\begin{aligned} \sigma_1^2 - \sigma_2^2 &= \lim_{T\to\infty} \int_0^T y^2 d(G_1-G_2) \\ &= \lim_{T\to\infty} T^2[G_1(T)-G_2(T)] - 2\lim_{T\to\infty} \int_0^T y(G_1-G_2)dy \\ &= 0 - 2\int_0^\infty y(G_1-G_2)dy \leqslant 0 \end{aligned}$$

を得る．これは (2b. 2. 7) なら $G_1 \geqslant G_2$ であるからである．しかし逆の命題は成り立たない．もし $\sigma_1^2 \leqslant \sigma_2^2$ ならば少なくとも x のある値にたいしては不等式 (2b. 2. 7) が成り立つが，すべての x に値にたいして成り立つとは限らない．

2b. 3　条件つき期待値

2a. 8 では X_1 のもとでの X_2 の条件つき分布関数をある特別な場合について定義した．

今度はより一般的な概念を条件つき期待値を定義することによって条件つき分布を表に出さずに導きたい．これによって条件つき確率が条件つき期待値の特別な場合であることがわかるであろう．

確率変数の対 (Y, T) を考える．ここでTは一般のn次元でもよい．Tを与えたときの Y の条件つき期待値 $E(Y|t)$ は次の性質をもつTの関数として定義される．すなわち，すべての $S \in \mathscr{B}_n$ にたいして

$$\int_{R_1 \times S} y dF(y, t) = \int_S E(Y|t) dF(t) \tag{2b. 3. 1}$$

ここで $R_1 \times S$ はT平面でSを底とする (Y, T) 平面での柱状集合である．このような $E(Y|t)$ の存在は次のようにしてわかる．(2b. 3. 1) の左辺は $P(S)$ について絶対連続である集合関数 $Q(S)$ であって $P(S)$ は

$$P(S) = \int_S dF(t) \tag{2b. 3. 2}$$

であらわされ，$P(S) = 0$ ならば $Q(S) = 0$ である．さらに $E(Y)$ が存在するならば，$Q(S)$ は可算加法的である．それらの条件のもとで，ラドン＝ニコディムの定理（付録2B参照）からほとんどいたるところ $[dF(t)]$ で定義されるTの関数gが存在し

$$Q(S) = \int_S g(t) dP = \int_S g(t) dF(t) \tag{2b. 3. 3}$$

であることがわかる．この関数 $g(t)$ を $E(Y|t)$ とすればよい．次に数理統計の重要な応用である条件つき期待値のいくつかの一般的性質を考えよう．まず，

（i）$E(Y) = \int E(Y|t) dF(t) = \underset{T}{E}[E(Y|T)]$　(2b. 3. 4)

ここでEの下の記号のTはTの分布関数に関して期待値を求めることを意味する．

さらに T に関して期待値をとるときは，$E(Y|t) = E(Y|T = t)$ は $E(Y|T)$ でおきかえられることを注意しておく．全体を通してこの約束に従うことにする．

もしSを全平面とすれば (2b. 3. 4) は (2b. 3. 1) からすぐ出てくる．確率密度が存在し，条件つき期待値が条件つき分布（密度）に関する期待値で定義されるときはその証明は容易であろう．

確率変数YとY^2を考え，$T = t$ のもとでYの条件つき分散を

$$V(Y|t) = E(Y^2|t) - [E(Y|t)]^2 \tag{2b. 3. 5}$$

と定義する．

（ii）$V(Y) = \underset{T}{E}[V(Y|T)] + \underset{T}{V}[E(Y|T)]$.

(2b. 3. 5) の両辺を $dF(t)$ について積分すると

$$\underset{T}{E}[V(Y|T)] = E(Y^2) - \underset{T}{E}[E(Y|T)]^2$$

$$= V(Y) - \underset{T}{V}[E(Y|T)]$$

である．このことから

$$V(Y) = \underset{T}{E}[V(Y|T)] + \underset{T}{V}[E(Y|T)] \tag{2b. 3. 6}$$

となりYの全分散を (a) 条件つき分散の平均と，(b) 条件つき平均の分散，という2つの成分の和としてあらわしている．

（iii）$g(T)$ をTのボレル関数で $E|Y| < \infty$，$E|Yg(T)| < \infty$ とすると次の2つの式が成り立つ：

（a）$E[g(T)Y|t] = g(t)E(Y|t)$　(2b. 3. 7)

（b）$E\{g(T)[Y - E(Y|T)]\} = 0.$　(2b. 3. 8)

gがTのボレル関数であるという仮定から，$g(T)$ 自身がTにより導かれる確率測度に関して確率変数であることが保証され，したがって期待値などが求められる．(2b. 3. 8) の結果は (2b. 3. 7) からの自明な帰結である．(2b. 3. 7) を証明するには T 平面における集合 $A \in \mathscr{B}_n$ の定義関数 $C_A(t)$ という T の簡単な関数を考える．こうして任意の集合 $S \in \mathscr{B}_n$ にたいして

$$\int_{R_1 \times S} y C_A(t) dF(y, t) = \int_S C_A(t) E(Y|t) dF(t) \tag{2b. 3. 9}$$

であり，(2b. 3. 7) は $C_A(t)$ について成立する．(2b. 3. 9) はgが単関数の有限の1次結合であるとき成立する．一般に，Tの有界集合上の簡単な関数の有限1次結合によってgを近似し，また条件 $E|g(T)Y| < \infty$ を使うことによって (2b. 3. 7) が導かれる．

$T = t$ のもとで集合 $A \in \mathscr{B}$ の条件つき確率は，確率変数 $Y(\cdot)$ を集合Aの定義関数として選ぶことによって条件つき期待値として示すことができる．$P(A|t) = E(Y|t)$ であることは条件つき確率の定義から証明される．ラドン＝ニコディムの定理は，Aを固定したときに測度 $[dF(t)]$ に関してほとんどいたるところでtの関数としての $P(A|t)$ が，存在することを保証しているだけであり，tの除外集合はAに依存することに注意しなければならない．したがって，確率1のt集合上ですべての A について $P(A|t)$ を定義することは――除外集合の和が確率0でない限り――可能でないであろうから，$T = t$ のときの $(\Omega, \mathscr{B})$ 上の条件つき確率測度はいつも存在するとはいえない．

しかし Kolmogorov (1933) と Doob (1953) は，まず $(\Omega, \mathscr{B})$ 上の条件つき確率測度をつくるのではなく，条件つき確率だけを使って確率変数Yの条件つき期待値を評価する方法を示唆している．かれらは，和

$$S(\lambda) = \sum_{-\infty}^{\infty} (j+1)\lambda P[j\lambda < Y(\omega) < (j+1)\lambda | t] \tag{2b. 3. 10}$$

を考えることによってそれを行った．この式は $E(Y)$ が存在するとき λ のもとで絶対収束であり，したがって

$$\lim_{\lambda\to 0} S(\lambda) = E(Y|t)$$

である．しかし，条件つき確率測度がどんなときに存在するかを確かめることはかなり興味あることである．

離散標本空間　もし標本空間に可算個の点だけがあるならば，条件つき確率測度は，いつでも定義される．任意の集合 $A \in \mathscr{B}$ の条件つき確率は，分母が 0 でないかぎり $P(A \cap \{T=t\})/P(T=t)$ で定義される．

ユークリッド標本空間　標本空間を R^n とするならば条件つき確率測度が存在する．その証明は $(R^n, \mathscr{B}_n)$ 上の確率測度は可算個の集合上でのみ，たとえば，有理頂点の座標が有理数である直方体の全体で定義されればよいということに基づく．このとき除外集合の和は確率測度ゼロである．長方形の集合上で得られた確率測度は $\mathscr{B}$ 内のすべての集合に一意的に拡張される．詳細については Doob (1953) を参照されたい．

標本空間の条件つき密度　多くの応用では，標本空間は R^n（n 次ユークリッド空間）であり，任意の点 $\mathbf{x} = (x_1, \cdots, x_n)$ で確率密度が存在する．もし $p(\mathbf{x})$ で密度をあらわすとするなら，集合 $A \in \mathscr{B}_n$ の確率は

$$P(A) = \int_A p(\mathbf{x}) dx_1 \cdots dx_n$$

で与えられる．いま T の可測関数を考え，これは多次元で確率密度 $q(t)$ をもつものとする．このとき $T=t$ の面上で $p(\mathbf{x})/q(t)$ という確率密度を定義できる．$\mathbf{x}$ の関数 $p(\mathbf{x})/q(t)$ は $T=t$ のもとでの R^n の点 $\mathbf{x}$ の条件つき密度とよばれる．もし T の各変数についての偏導関数が連続ならば，多重積分の知識から体積元は

$$dx_1 \cdots dx_n = dt dS \tag{2b. 3. 11}$$

のようにあらわされる．ここで dS は面分である．このとき t に関する積分は，はじめに dS で積分する——t を固定して——次に dt で積分する．この性質から，この節で考えられるすべての結論を十分理解できよう．もし Y が X の任意の測度関数ならば $T=t$ のときその条件つき期待値は

$$E(Y|t) = \int Y \frac{p(\mathbf{x})}{q(t)} dS \tag{2b. 3. 12}$$

で，これは t の関数である．

2b. 4　特　性　関　数

分布関数 F をもつ確率変数 X の特性関数 $f(t)$ の定義は（積分範囲は特に断わらないかぎり $(-\infty, \infty)$ である）

$$f(t) = \int e^{itx} dF(x)$$

$$= \int \cos tx\, dF(x) + i \int \sin tx\, dF(x) \tag{2b. 4. 1}$$

である．ここで $i = \sqrt{-1}$ である．(2b. 4. 1) の右辺の積分は，すべての x にたいして $|e^{itx}| = |\cos tx + i \sin tx| = 1$ で有界だから必ず存在する．解析学の標準的な結果を使って，次の諸結果が証明できる．

（i）$f(0) = 1$，$|f(t)| \leqslant 1$，$f(-t) = \overline{f(t)}$，$f(t)$ の共役複素数．

（ii）$f(t)$ は全実数軸上で一様連続である．

$\cos tx$ と $\sin tx$ は全実数軸上で連続でまた有界ということから上の結果が得られる．

（iii）$\overline{f(t)}$ は $-X$ の特性関数である．$f(t)$ が実数なのは，X が対称であるときで，そのときにかぎる．$bX + a$ の特性関数は $f(bt)\exp(ita)$ である．

（iv）もし $\alpha_r = E(X^r)$ が存在するならば，$f(t)$ の r 次導関数

$$f^{(r)}(t) = (i)^r \int x^r e^{itx} dF(x) \tag{2b. 4. 2}$$

が存在し，一様連続である．また

$$\alpha_s = E(X^s) = (i)^{-s} f^s(0), \qquad s \leqslant r \tag{2b. 4. 3}$$

である．

$|x^r e^{itx}| \leqslant |x^r|$ で，$E(X^r)$ が存在することから (2b. 4. 2) の右辺の積分は絶対収束で，その場合，微分と積分とは交換可能である（付録 **2B** の (2B. 14) 参照）．このことから

$$\frac{d^r f}{dt^r} = \int \frac{d^r e^{itx}}{dt^r} dF(x) = (i)^r \int x^r e^{itx} dF(x)$$

が得られる．$f^{(r)}$ の連続性は e^{itx} の連続性と有界性とから示される．(2b. 4. 3) は容易に導ける．

（v）もし $f^{(2n)}(0)$ が存在し，有限であるならば，$r \leqslant 2n$ にたいして $E(X^r) < \infty$．

なぜなら

$$|f^{(2n)}(0)| = \left| \lim_{u\to 0} \int \left(\frac{e^{iux} - e^{-iux}}{2u} \right)^{2n} dF(x) \right|$$

$$= \lim_{u\to 0} \int \left(\frac{\sin ux}{ux} \right)^{2n} x^{2n} dF(x) \geqslant \int x^{2n} dF(x). \tag{2b. 4. 4}$$

ここで極限と積分の交換に関してファトゥーの補助定理（付録 **2B** の (2B. 5) 参照）を使っている．そこで (v) は不等式 (2b. 4. 4) から導かれ，また $E(X^s)$ が有限であることは不等式 $|x^k| \leqslant 1 + |x|^s$ を考慮した場合 $k \leqslant s$ なる k にたいして $E(X^k)$ が有限であることを意味している．

（vi）$E(X^r)$ が存在すれば，$f(t)$ は $t \to 0$ で次のようにテイラー展開される．

$$f(t) = \sum_0^r \alpha_j \frac{(it)^j}{j!} + o(t^r). \tag{2b. 4. 5}$$

（vii）$\phi(t)=\log f(t)$ とするとき，これを X（またはその分布関数F）の第2特性関数という．もし $E(X^r)$ が存在するならば $\phi(t)$ は $t\to 0$ でテイラー展開され，

$$\phi(t)=\sum_0^r \kappa_j\frac{(it)^j}{j!}+o(t^r) \tag{2b. 4. 6}$$

を得る．

$\phi(t)$ の展開式の係数 κ_j をキュムラントまたは半不変量という（半不変量と呼ばれる理由は—— κ_1 のときを除いて——，それが確率変数の平行移動にたいして不変量であるからである）．いま確率変数 $X+a$ を考えるとその特性関数は $e^{ita}f(t)$ で，第2特性関数は

$$\log e^{ita}f(t)=ita+\log f(t)=ita+\phi(t)$$

と書け，t の係数が変わるだけである．ところが，この性質は，積率係数 α_j にたいしては成り立たない．

（vi）と（vii）との展開から，κ_j，α_j と中心積率 μ_j との結合関係が得られる．

$$\begin{aligned}
\kappa_1 &= \alpha_1 &&= \mu\\
\kappa_2 &= \alpha_2-\alpha_1^2 &&= \mu_2\\
\kappa_3 &= \alpha_3-3\alpha_1\alpha_2+2\alpha_1^3 &&= \mu_3\\
\kappa_4 &= \alpha_4-3\alpha_2^2-4\alpha_1\alpha_3+12\alpha_1^2\alpha_2-6\alpha_1^4 &&= \mu_4-3\mu_2^2
\end{aligned}$$

また，逆に次のようにもあらわせる．

$$\begin{aligned}
\alpha_1&=\kappa_1, & \alpha_3&=\kappa_3+3\kappa_2\kappa_1+\kappa_1^3\\
\alpha_2&=\kappa_2+\kappa_1^2, & \alpha_4&=\kappa_4+3\kappa_2^2+4\kappa_1\kappa_3+6\kappa_1^2\kappa_2+\kappa_1^4.
\end{aligned}$$

また，係数

$$\sqrt{\beta_1}=\frac{\mu_3}{\mu_2^{3/2}}=\frac{\kappa_3}{\kappa_2^{3/2}}$$

$$\beta_2=\frac{\mu_4}{\mu_2^2}=\frac{\kappa_4}{\kappa_2^2}+3$$

をそれぞれ歪みの測度，尖りの測度という．

積率母関数　関数

$$M(t)=\int e^{tx}dF(x) \tag{2b. 4. 7}$$

がもし実数直線上の，t の原点を含む区間で存在するならば，その関数は積率母関数とよばれている．

確率母関数　関数 $f(t)$，$\phi(t)$，$M(t)$ は任意の確率変数にたいして定義される．ただし，$M(t)$ はすべての場合については存在しないかもしれない．値 $0,1,\cdots$ をとるような離散確率変数Xにたいしては，別の関数

$$P(t)=\sum_0^\infty p_rt^r,\qquad -1\leqslant t\leqslant 1$$

が定義される．この関数を確率母関数という．確率母関数はまたXの積率の決定にも有用である．この場合には次の関係

$$\left.\frac{d^sP}{dt^s}\right|_{t=1}=E[X(X-1)\cdots(X-s+1)]$$

を用いる．ここで，右辺はXの s 次階乗積率といい，ϕ_s であらわす．階乗積率はもとの積率 α_s から次の式によって求められる．その式は

$$\phi_1=\alpha_1,\quad \phi_2=\alpha_2-\alpha_1,\quad \phi_3=\alpha_3-3\alpha_2+2\alpha_1,\quad \phi_4=\alpha_4-6\alpha_3+11\alpha_2-6\alpha_1$$

または逆に

$$\alpha_1=\phi_1,\quad \alpha_2=\phi_2+\phi_1,\quad \alpha_3=\phi_3+3\phi_2+\phi_1,\quad \alpha_4=\phi_4+6\phi_3+7\phi_2+\phi_1$$

である．

ある特定の確率変数の確率母関数，積率母関数，特性関数 $[P(t),\ M(t),\ f(t)]$ を導いておこう．それらの関数はあとで使う．

2 項分布

$$p_r=\binom{n}{r}\pi^r(1-\pi)^{n-r},\qquad r=0,1,\cdots,n$$

$$P(t)=\sum_1^n\binom{n}{r}\pi^r(1-\pi)^{n-r}t^r=(1-\pi+\pi t)^n,$$

$$M(t)=P(e^t)=(1-\pi+\pi e^t)^n,$$

$$f(t)=M(it)=(1-\pi+\pi e^{it})^n. \tag{2b. 4. 8}$$

ポアソン分布

$$p_r=e^{-\lambda}\frac{\lambda^r}{r!},\qquad r=0,1,\cdots,\infty$$

$$P(t)=\sum e^{-\lambda}\frac{\lambda^rt^r}{r!}=e^{-\lambda(1-t)},$$

$$M(t)=P(e^t)=e^{-\lambda(1-e^t)}$$

$$f(t)=M(it)=e^{-\lambda(1-e^{it})}. \tag{2b. 4. 9}$$

正規分布

$$p(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-(x-\mu)^2/2\sigma^2}$$

$$\begin{aligned}
M(t)&=\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{tx}e^{-(x-\mu)^2/2\sigma^2}dx\\
&=e^{t\mu+t^2\sigma^2/2}\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-(y-t\sigma^2)^2/2\sigma^2}dy,\qquad y=x-\mu\\
&=e^{t\mu+t^2\sigma^2/2}. \qquad\qquad (2b. 4. 10)
\end{aligned}$$

$$f(t)=M(it)=e^{it\mu-t^2\sigma^2/2}. \tag{2b. 4. 11}$$

分布関数Fは $F(a+0)-F(a-0)=1$ となるRのただ1つの増加点aをもつとき退

化とよばれる．その場合Fの特性関数は $f(t)=\exp(ita)$ である．

(viii)　fが退化分布関数の特性関数であるのは，αを無理数として$t\neq 0$ および $\alpha t\neq 0$ の2つの値にたいして，関数fの絶対値が1になるときである．特に，もし非退化区間で $|f(t)|=1$ ならばfは退化である．

(ix)　**分布関数のたたみこみ**　R上の関数Fは

$$F(x)=\int F_1(x-y)dF_2(y),\qquad x\in R \tag{2b.4.12}$$

が成り立つとき分布関数 F_1 と F_2 とのたたみこみとよばれ $F=F_1*F_2$ であらわす．

(a)　(2b.4.12) で定義されたFは分布関数で，もし f, f_1, f_2 が F, F_1, F_2 の各々の特性関数であるならば $f=f_1f_2$ である．

F_1 が非減少関数であるときはFもまた非減少関数である．$F_1(x-y)\leqq 1$ であるからルベーグの上限収束定理（付録 **2B** の (2B.11) 参照）により $F(-\infty)=0$, $F(\infty)=1$ である．さらに F_1 が左側連続だからルベーグの上限収束定理より関数Fの左側連続であることがわかる．これで最初の命題は証明された．

第2の命題を証明しよう．　いま $a<x_{n1}<\cdots<x_{n,k_n+1}=b$ で $n\to\infty$ のとき $\sup_k(x_{n,k+1}-x_{nk})\to 0$ であるとする．このときすべての $t\in R$ にたいして次式を得る：

$$\int_a^b e^{itx}dF(x)=\lim\sum_k e^{itx_{nk}}F[x_{nk},x_{n,k+1})$$

$$=\lim\int\sum_k e^{it(x_{nk}-y)}F_1[x_{nk}-y,x_{n,k+1}-y)e^{ity}dF_2(y).$$

このことから

$$\int_a^b e^{itx}dF(x)=\int\left\{\int_{a-y}^{b-y}e^{itx}dF_1(x)\right\}e^{ity}dF_2(y)$$

である．ここで $F[c,d)=F(d)-F(c)$ である．$a\to\infty$, $b\to\infty$ とすると

$$\int e^{itx}dF(x)=\int e^{itx}dF_1(x)\int e^{ity}dF_2(y)$$

だから $f=f_1f_2$ である．

もし $f=f_1f_2$ とすると特性関数の一意性（**2b.5** 参照）から $F=F_1*F_2$ であることに注意しよう．特性関数の積もまた特性関数であること，また特に，f が特性関数ならば $|f|^2$ もまた特性関数であることがわかる．

(b)　XとYとが互いに独立な確率変数で，それぞれ特性関数が f_1, f_2 であり，fが $T=X+Y$ となるTの特性関数であるとすると

$$f=E\{e^{it(X+Y)}\}=E(e^{itX})\cdot E(e^{itY})=f_1f_2 \tag{2b.4.13}$$

が成り立つ．

(a)の結果を使うとたたみこみ F_1*F_2 は分布関数 F_1, F_2 をもつ互いに独立な分布関数の和の分布関数であることがわかる．$F=F_1*F_2$ は F_1 と F_2 の合成ともよばれる．

2b.5　反　転　定　理

分布関数を与えると，(2b.4.1) で定義される特性関数が存在する．すると特性関数から分布関数が一意的にきまるかどうかという問題が自然に生じる．一意的にきまるというのが答えで，その正確な結果は (i) で述べる．

分布関数があらゆる次数の積率 $\alpha_1, \alpha_2, \cdots$ をもつとしよう．与えられた積率系列は分布関数を一意的に定義するであろうか？　この結果は一般には成り立たない．一意性の十分条件は (ii) で与えられる．

(i)　$f(t)$ が $F(x)$ に対応する特性関数で，$F(x)$ は区間 $(a-c, a+c)$ の端点で連続であるとすると

$$F(a+c)-F(a-c)=\lim_{T\to\infty}\frac{1}{\pi}\int_{-T}^{T}\frac{\sin ct}{t}e^{-ita}f(t)dt, \tag{2b.5.1}$$

である．それゆえ，もし2つの分布関数が存在するならば，両方の分布のすべての連続な点でそれらは一致し，したがって，まったく同一のものになる．

いま，式

$$J=\frac{1}{\pi}\int_{-T}^{T}\frac{\sin ct}{t}e^{-ita}f(t)dt$$

$$=\frac{1}{\pi}\int_{-T}^{T}\frac{\sin ct}{t}e^{-ita}dt\int e^{itx}dF(x)$$

を考える．$|(\sin ct/t)e^{it(x-a)}|<|c|$ であるから，積分の順序は交換可能であり，J は次のようにあらわせる：

$$\frac{1}{\pi}\int dF(x)\int_{-T}^{T}\frac{\sin ct}{t}e^{-it(a-x)}dt=\int_{-\infty}^{\infty}g(T,x)dF(x)$$

ただし，$g(T,x)$ は

$$\frac{1}{\pi}\int_{-T}^{T}\frac{\sin ct}{t}e^{-it(a-x)}dt=\frac{2}{\pi}\int_0^T\frac{\sin ct}{t}\cos(x-a)t\,dt$$

に等しい．ここで，公式 $2\sin A\cos B=\sin(A+B)+\sin(A-B)$ を利用することによって次式を得る：

$$g(T,x)=\frac{1}{\pi}\int_0^T\frac{\sin(x-a+c)t}{t}dt+\frac{1}{\pi}\int_0^T\frac{\sin(c+a-x)t}{t}dt \tag{2b.5.2}$$

$$\lim_{T\to\infty}g(T,x)=\frac{1}{\pi}\int_0^\infty\frac{\sin(x-a+c)t}{t}dt+\frac{1}{\pi}\int_0^\infty\frac{\sin(c+a-x)t}{t}dt.$$

次に式 (2b. 5. 2) の各項にたいして次の結果

$$\frac{1}{\pi}\int_0^\infty \frac{\sin \alpha t}{t}dt = \begin{cases} \frac{1}{2} & \alpha > 0 \text{ のとき} \\ 0 & \alpha = 0 \text{ のとき} \\ -\frac{1}{2} & \alpha < 0 \text{ のとき} \end{cases}$$

を使う．すると，

$$\lim_{T\to\infty} g(T, x) = \begin{cases} 0 & x < a - c \text{ または } x > (a + c) \text{ のとき} \\ \frac{1}{2} & x = a - c \text{ または } x = a + c \text{ のとき} \\ 1 & a - c < x < a + c \text{ のとき} \end{cases}$$

を得る．なぜなら，上の式の2つの項の値は，第1の場合は $-\frac{1}{2}$, $\frac{1}{2}$ または $\frac{1}{2}$, $-\frac{1}{2}$ となり，第2の場合は 0, $\frac{1}{2}$ または $\frac{1}{2}$, 0 となり，第3の場合は $\frac{1}{2}$, $\frac{1}{2}$ となるからである．こうして $\lim g(T, X)$ の存在が証明できた．

さらに $|g(T, x)|$ は，$\int_0^T (\sin bt/t)dt$ が T に関して一様に有界であることから，有界である．このことから，ルベーグの上限収束定理（付録 **2B**）によって

$$\lim_{T\to\infty}\int_{-\infty}^{\infty} g(T, x)dF(x) = \int_{-\infty}^{\infty} \lim_{T\to\infty} g(T, x)dF(x)$$
$$= \int_{a-c}^{a+c} dF(x) = F(a + c) - F(a - c)$$

であることがわかる．これで (2b. 5. 1) が証明された．

特に，$|f(t)|$ が積分可能であるときには，(2b. 5. 1) の両辺を c で割って c を 0 に近づけると，

$$\lim_{c\to 0}\frac{F(a + c) - F(a - c)}{2c} = \frac{1}{2\pi}\int_{-\infty}^{\infty} \lim_{c\to 0} \frac{\sin ct}{ct} e^{-ita} f(t)dt \qquad (2b. 5. 3)$$

$$F'(a) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-ita} f(t)dt \qquad (2b. 5. 4)$$

を得る．したがって $F(x)$ の導関数が存在し，それはまた連続である．(2b. 5. 3) の被積分関数の絶対値は $|f(t)|$ より小さく，$|f(t)|$ の積分可能性から上限収束定理により積分記号の内側に極限記号をもってきてよいことを注意しておこう．この目的のためには $|f(t)|$ の積分可能性が強い条件となっていると思われる．

（ii）**積率問題**　$\alpha_1, \alpha_2, \cdots$ を確率変数の積率系列とする．その積率系列は一般には分布関数を一意的にきめない．分布関数が一意的にきまるための十分条件は，級数：

$$\sum_1^\infty \frac{\alpha_i}{i!} t^i \qquad (2b. 5. 5)$$

が，ある $t > 0$ にたいして絶対収束することである．

スティルチェスの積率問題として知られているこの命題の証明の本質は，条件 (2b. 5. 5) のもとで特性関数が一意的に定まることを示すことにある．証明の詳細については Hamburger (1920, 1921) の原論文を参照していただきたい．2次元確率変数の場合の十分条件は Cramer and Wold (1936) に与えられている．

(2b. 5. 5) が必要条件でないことは次のことからわかるであろう．すなわち，X と Y とが互いに独立な変数で密度がそれぞれ e^{-x}, e^{-y} であるとき，$Z = X \log(1 + Y)$ の積率から以下に示す (b) によって Z の分布が一意に定まる．しかし，条件 (2b. 5. 5) はみたされていないのである (Dharmadhikari (1965) 参照)．

以下に示すどの1つも積率系列 $\alpha_1, \alpha_2, \cdots$ から分布関数が一意に定まるための十分条件である．

（a）確率変数の範囲は有限である．

（b）確率変数の範囲が $(-\infty, \infty)$ であるとき，$\sum_0^\infty (\alpha_{2j})^{-1/2j} = \infty$ である．

（c）確率変数の範囲が $(0, \infty)$ であるとき，$\sum_0^\infty (\alpha_j)^{-1/2j} = \infty$ である．

（d）$\limsup_{n\to\infty}[(\alpha_{2n})^{1/2n}/2n]$ は有限である．

2b. 6　多変量積率

$F(x_1, \cdots, x_k)$ が確率変数 $X_1, \cdots, X_k$ の同時分布関数であるとするとき，定数 $a_1, \cdots, a_k$ のまわりの多変量積率（モーメント）$\alpha_{r_1\cdots r_k}$ は

$$E[(X_1 - a_1)^{r_1}\cdots(X_k - a_k)^{r_k}] = \int\cdots\int (x_1 - a_1)^{r_1}\cdots(x_k - a_k)^{r_k} dF(x_1, \cdots, x_k)$$

である．$a_i = E(X_i) = \mu_i$ のときは，1変数の場合と同様に中心積率となる．2次積率

$$E(X_i - \mu_i)(X_j - \mu_j)$$

を X_i と X_j との共分散とよび，$\mathrm{cov}(X_i, X_j)$ または単に $C(X_i, X_j)$ であらわす．$C(X_i, X_j)$ を σ_{ij} と書くと次の結果が得られる：

1. $\sigma_{ii} = C(X_i, X_i) = \sigma_i^2$　（X_i の分散）
2. $\sigma_{ij}^2 \leqslant \sigma_{ii}\sigma_{jj}$　（コーシー＝シュヴァルツの不等式の適用による）
3. $-1 \leqslant \rho_{ij} \leqslant 1$，ここで $\rho_{ij} = \sigma_{ij}/\sqrt{\sigma_{ii}\sigma_{jj}}$ を X_i と Y_j との積率相関係数とよぶ［3は2から導かれる］．
4. $a_1, \cdots, a_k$ を定数とするとき，$a_1X_1 + \cdots + a_kX_k$ の分散 $V(a_1X_1 + \cdots + a_kX_k)$ は

$$E[\sum a_i(X_i - \mu_i)]^2 = \sum\sum a_ia_j\sigma_{ij} \qquad (2b. 6. 1)$$

と書ける．同様に，$\sum a_i X_i$ と $\sum b_i X_i$ との共分散は

$$\sum\sum a_i b_j \sigma_{ij} \tag{2b. 6. 2}$$

であらわせる．行列記号で，$\mathbf{\Sigma} = (\sigma_{ij})$ とし，$\mathbf{A}, \mathbf{B}, \mathbf{X}, \mathbf{M}$ をそれぞれ係数 a_i, b_i，確率変数 X_i，平均値 μ_i の列ベクトルとすると，

$$\begin{aligned} &E(\mathbf{A}'\mathbf{X}) = \mathbf{A}'E(\mathbf{X}) = \mathbf{A}'\mathbf{M} \\ &V(\mathbf{A}'\mathbf{X}) = \mathbf{A}'\mathbf{\Sigma}\mathbf{A}, \qquad V(\mathbf{B}'\mathbf{X}) = \mathbf{B}'\mathbf{\Sigma}\mathbf{B} \\ &C(\mathbf{A}'\mathbf{X}, \mathbf{B}'\mathbf{X}) = \mathbf{A}'\mathbf{\Sigma}\mathbf{B} = \mathbf{B}'\mathbf{\Sigma}\mathbf{A} \end{aligned} \tag{2b. 6. 3}$$

である．ここでVは分散，Cは共分散をあらわす．

5.　行列 $\mathbf{\Sigma}$ は，2次形式 $\sum\sum a_i a_j \sigma_{ij}$ が非負で確率変数の分散をあらわすことから，非負値である（(2b. 6. 1) 参照）．

6.　$\mathbf{\Sigma}$ は，確率変数 $X_1, \cdots, X_k$ の間に（確率1で）線形関係が存在しないとき，かつ，そのときに限り，正値である．

7.　k 次元ベクトル確率変数Xの分散共分散行列 $\mathbf{\Sigma}$（非負定符号）の階数は $r < k$ であるとする．このとき $\mathbf{X}$ は

$$\mathbf{X} = \mathbf{B}\mathbf{Y} + \mathbf{C} \tag{2b. 6. 4}$$

であらわされる．ここで $\mathbf{Y}$ は分散共分散行列 $\mathbf{I}$ をもつ r 次元ベクトル確率変数，$\mathbf{B}$ は階数 r の $k \times r$ 行列で，$\mathbf{C}$ は定数の k 次元ベクトルである．

多変量問題においては，変数の部分集合の条件つき分布，たとえば $X_{r+1}, \cdots, X_k$ が与えられたときの $X_1, \cdots, X_r$ の分布，がよく研究される．特に他の変数が与えられたときの，ある変数の条件つき期待値と条件つき分散が，重要な役割を演ずる．これらの条件つき積率やその計算に関する問題の細かい吟味は第4章の **4g. 1**，**4g. 2** で行う．

2c　極　限　定　理

2c. 1　コルモゴロフの一致性定理

$(\Omega, \mathscr{B}, P)$ を基本確率空間，$\{X_i(\cdot)\}$，$i = 1, 2, \cdots$ を確率変数の無限列とし，F_n を R^n で定義された $(X_1(\cdot), \cdots, X_n(\cdot))$ の分布関数とする．このとき次の一致性 $F_n(t_1, \cdots, t_n) = F_{n+1}(t_1, \cdots, t_n, \infty)$ がある．

いま，逆の命題を考えよう．すなわち，R^n，$n = 1, 2, \cdots$ 上の分布関数 F_n の列が与えられ一致性をもつならば，$(\Omega, \mathscr{B}, P)$ が存在し，また F_n が $(X_1(\cdot), \cdots, X_n(\cdot))$，$n = 1, 2, \cdots$ の分布関数であるような確率変数 $X_i(\cdot)$ の列が存在するということである．これから証明するコルモゴロフの定理はこのことをいっている．

実際，Ωとして無限数列を要素とするユークリッド空間 R^∞ を選び，i 番目の確率変数として関数の値がその点の i 番目の座標であるようなものを選び，$\mathscr{B}$ と P を適当に定義する．

確率変数 $X_{i_1}, \cdots, X_{i_n}$ の有限次元部分標本空間をとり，そのなかにボレル集合 A_{i_n} を考えよう．A_{i_n} の各点 $(x_{i_1}, \cdots, x_{i_n})$ に他の座標の任意の値をつけ加えて無限数列をつくり，その全体を A_i とする．このようにして得られた R^∞ の部分集合を，ある E_n 内の A_{i_n} を底にもつボレル集合という．$P_n(A_{i_n})$ を，**2a. 6** で考えたように，$X_{i_1}, \cdots, X_{i_n}$ の分布関数によって適当な E_n 上の $\mathscr{B}_n$ に導入される集合 A_{i_n} の確率とする．いま，

$$P_\infty(A_i) = P_n(A_{i_n}) \tag{2c. 1. 1}$$

と定義する．このことから，関数 P_∞ はすべてのボレル柱集合にたいして定義される．ボレル柱集合は確率変数 $X_1, X_2, \cdots$ のすべての有限部分集合を考えることによって与えられる．ボレル柱集合の族は，〔ブール〕集合体 $\mathscr{F}_0$ をなす．(2c. 1. 1) で定義された P_∞ は $\mathscr{F}_0$ 上の確率測度を構成する．

P_∞ が確率測度であるためには P_∞ の可算加法性がみたされなければならない．それは，もし，空集合 $\varnothing$ に近づくような $\mathscr{F}_0$ 内の集合 $A_1, A_2, \cdots$ の減少列があるとすると $i \to \infty$ のとき $P_\infty(A_i) \to 0$ である，と同値である．

一般性を失わずに次のことを仮定することができる．すなわち A_1 の底が X_1 軸上の集合 A_{11} で，A_2 の底が (X_1, X_2) 空間内の集合 A_{22}，$\cdots$，A_n の底が $(X_1, \cdots, X_n)$ 空間内の集合 A_{nn} などであるということである．A_{nn} を与えると，$(X_1, \cdots, X_n)$ 空間内に有界な閉集合 B_{nn} が存在し，B_{nn} は $B_{nn} \subset A_{nn}$ で $P_n(A_{nn} - B_{nn}) < \delta/2^n$ である．いま，B_n を底 B_{nn} をもつ柱集合とする．このとき $B_n \subset A_n$ であり，また

$$P_\infty(A_n - B_n) = P_n(A_{nn} - B_{nn}) < \frac{\delta}{2^n}$$

が成り立つ．$A_n = \bigcap_1^n A_j$ および $C_n = \bigcap_1^n B_j$ を考えよう．このとき，

$$P_\infty(A_n - C_n) \leqslant \sum_1^n P_\infty(A_i - B_i) < \delta\left(\frac{1}{2} + \cdots + \frac{1}{2^n}\right) < \delta \tag{2c. 1. 2}$$

である．

さて，すべての n にたいして $P_\infty(A_n) > \varepsilon$ であって，0に近づかないものと仮定しよう．$\delta = \varepsilon/2$ を選び出すと，(2c. 1. 2) から

$$P_\infty(C_n) > P_\infty(A_n) - \delta > \varepsilon/2 > 0$$

を得る．これは集合 C_n がすべての n にたいして空集合ではないことを意味する．$\xi_n = (\xi_{n1}, \xi_{n2}, \cdots)$ を C_n に属する R^∞ の点とする．$C_n = \bigcap_1^n B_i$ であるからすべての n にた

いして $\xi_n \in B_1$ である．点 $\xi_1, \xi_2, \cdots$ の第1座標は B_{11} に属する．そして B_{11} は閉集合で有界であることから点の部分列が存在し，その第1座標は極限たとえば b_1 に近づく．点 $\xi_2, \xi_3, \cdots$ は B_2 に属する．したがって点 $\xi_2, \xi_3, \cdots$ の最初の2つの座標は B_{22} に属する．最初の部分列の点のうちで $\xi_2, \xi_3, \cdots$ に属するようなものだけを考えよう．B_{22} が有界で閉集合であることから第2座標の極限が b_2 になり，第1座標の極限が b_1 になるような最初の部分列の部分列が存在する．以下同様である．次に点 $b = (b_1, b_2, \cdots)$ を考えよう．いま，点 $(b_1, \cdots, b_k)$ が B_{kk} に属すから，すべての k にたいして $b \in B_k$ である．このことから $b \in \bigcap B_k = \bigcap C_k$ である．すなわち $\bigcap C_k \neq \varnothing$ である．したがって，

$$C_k \subset A_k \Rightarrow \bigcap C_k \subset \bigcap A_k \Rightarrow \lim A_k = \bigcap A_k \neq \varnothing$$

は矛盾である．このことから $P_\infty(A_i) \to 0$ となる．

関数 P_∞ は $\mathscr{F}_0$ のボレル拡張 $\mathscr{B}(\mathscr{F}_0)$（これを $\mathscr{B}_\infty$ であらわす）上の唯一の確率測度 Q を定義する．測度空間 $(R^\infty, \mathscr{B}_\infty, Q)$ はすべての有限次元の分布により完全に規定される．

特に，値 ω での関数値列のいくつかの極限関数の分布に興味をもつ．それらの関数は

$$\lim X_n(\omega), \quad \sup X_n(\omega), \quad \overline{\lim}\, X_n(\omega), \quad \underline{\lim}\, X_n(\omega)$$

などである．読者は，そのような関数が有界であれば，それもまた確率変数（すなわち可測関数）であることを理解できよう．

2c.2　確率変数列の収束

確率変数の列 $\{X_n\}$，$n = 1, 2, \cdots$ の定数 c への収束を次のように名づける．

（a）弱収束または確率収束（$X_n \xrightarrow{P} c$ と書く）　すべての与えられた $\varepsilon > 0$ にたいして

$$\lim_{n\to\infty} P(|X_n - c| > \varepsilon) = 0 \tag{2c.2.1}$$

であるとき．

（b）強収束またはほとんど確実な収束（確率1で $\lim_{n\to\infty} X_n = c$ または $X_n \xrightarrow{\text{a.s.}} c$ と書く）

$$P(\lim_{n\to\infty} X_n = c) = 1 \tag{2c.2.2}$$

または，これと同じことであるが，すべての ε にたいして

$$\lim_{N\to\infty} P(\sup_{n\geq N} |X_n - c| > \varepsilon) = 0 \tag{2c.2.3}$$

であるとき．

（c）2乗平均収束（$X_n \xrightarrow{\text{q.m.}} c$ と書く）

$$\lim_{n\to\infty} E(X_n - c)^2 = 0 \tag{2c.2.4}$$

であるとき．

定義（2c.2.1）と（2c.2.4）には難かしいことはない．これらの定義は有限次元の確率変数に関して確率や期待値を計算することと，変数の数が大きくなるときの極限をとることに依存しているからである．しかし，定義（2c.2.2）と（2c.2.3）に含まれる概念はいささか深遠である．この確率はその要素がある一定の性質を満足しているような R^∞ 内の集合に直接関係するのである．

確率変数の列 $\{X_n(\cdot)\}$，$n = 1, 2, \cdots$ が（a），（b），または（c）の意味である確率変数 $X(\cdot)$ に収束するとは，$\{X_n(\cdot) - X(\cdot)\}$，$n = 1, 2, \cdots$ が（a），（b），または（c）の意味で0に収束することをいう．

次の結果はいろいろな形の収束の間の関係を示す．

（i）(c) ⇒ (a)　<u>すなわち，2乗平均収束 ⇒ 確率収束．</u>

これはチェビシェフの不等式（2b.2.6）

$$P(|X_n - c| > \varepsilon) \leqslant \frac{1}{\varepsilon^2} E(X_n - c)^2$$

からただちにわかる．

（ii）(b) ⇒ (a)　<u>すなわち，ほとんど確実な収束 ⇒ 確率収束．</u>

$$|X_N - c| > \varepsilon \Rightarrow \sup_{n\geq N} |X_n - c| > \varepsilon$$

および

$$P(|X_n - c| > \varepsilon) \leqslant P(\sup_{n\geq N} |X_n - c| > \varepsilon)$$

から求める結果が導ける．ほかの条件がなければ，(a)，(b)，(c) の間にこれら以外の関係は成り立たない．

（iii）<u>$\sum_1^\infty E(X_n - c)^2 < \infty$ のもとで，$X_n \xrightarrow{\text{q.m.}} c$ であるならば，$X_n \xrightarrow{\text{a.s.}} c$ である．</u>

$\sum_1^\infty E(X_n - c)^2 < \infty$ であるから，無限和

$$\sum_1^\infty (X_n - c)^2 \tag{2c.2.5}$$

は R^∞ において測度零の系列の集合を除いては収束しなければならない．(2c.2.5) が測度 p の集合上で発散するものとしよう．A_N を $\sum_1^N (X_i - c)^2 > \lambda$ となるような R^∞ 内の集合とする．系列 $\{A_N\}$ は非減少で，発散列はいつかは任意に与えられた値 λ をこえるから，(2a.3.1) により

$$\lim P(A_N) = P(\lim A_N) \geqslant p$$

である．したがって

$$\sum_1^N E(X_i - c)^2 = E\Big[\sum_1^N (X_i - c)^2\Big] \geqslant \lambda p$$

となり，λを十分大きくとれば，

$$\sum_1^\infty E(X_i - c)^2 < \infty$$

であるという仮定に矛盾する．

このことから $P(\sum_1^\infty (X_n - c)^2$ は収束する$) = 1$ である．しかし，もし無限系列が収束するならば，$n \to \infty$ のときそのn番目の項は0に近づく．このことから．

$$P(|X_n - c| \to 0,\ n \to \infty \text{ のとき}) \geqslant P\Big(\sum_1^\infty (X_n - c)^2 \text{ は収束する}\Big) = 1$$

を得る．

これに付随して，Y_n，$n = 1, 2, \cdots$ を確率変数の列とするとき，

$$E\Big(\sum_1^\infty Y_n\Big) = \sum_1^\infty E(Y_n) \tag{2c. 2. 6}$$

が，$\sum_1^\infty E|Y_n| < \infty$ という条件のもとで成り立つ．これは $\sum_1^\infty Y_n$ が確率 1 で収束することを保証する．

2c. 3　大 数 の 法 則

ある未知の量（パラメータ）の推定は，実際には，その量の個々の測定値には誤差を伴うけれども，多数回繰り返し測定をして，その平均をとることによって行われる．それゆえ，そのような推定値の性質を調べることは興味がある．まず第1には，測定の個数がふえた（$\to\infty$）ときの推定値の行動についてである．すなわち，その推定値はいま問題にしているパラメータの真値になんらかの意味で収束するであろうか？　この問題は次のように定式化される．

$\{X_n\}$，$n = 1, 2, \cdots$ を観測値の系列，$\bar{X}_n$ を最初のn個の観測値の平均値とする．このとき，いかなる条件のもとで，**2c. 2** で考えた (a), (b), (c) の収束のうちのいずれかで，

$$\bar{X}_n \to \xi \quad (\text{未知の量}) \tag{2c. 3. 1}$$

と主張できるであろうか？　そのため問題をさらに一般化し，$\{\xi_n\}$，$n = 1, 2, \cdots$ を観測値の系列 $\{X_n\}$，$n = 1, 2, \cdots$ によって測定されるべき定数の系列とするとき，

$$\bar{X}_n - \bar{\xi}_n \to 0 \tag{2c. 3. 2}$$

が成立する条件を調べよう．(2c. 3. 1) や (2c. 3. 2) の収束が起きるとき，大数の法則が成り立つという．収束が (2c. 2. 1) の意味で“確率的”であるとき，大数の弱法則が成り立つといい，(2c. 2. 2) の意味で“確率 1”または“ほとんど確実な”とき，大数の強法則が成り立つという．

（i）**チェビシェフの定理（大数の弱法則）**　$E(X_i) = \mu_i$，$V(X_i) = \sigma_i^2$，$\mathrm{cov}(X_i, X_j) = 0$，$i \neq j$ とするとき

$$\lim_{n\to\infty} \frac{1}{n^2} \sum_1^n \sigma_i^2 = 0 \Rightarrow \bar{X}_n - \bar{\mu}_n \xrightarrow{P} 0 \tag{2c. 3. 3}$$

である．

チェビシェフの不等式により

$$P(|\bar{X}_n - \bar{\mu}_n| > \varepsilon) \leqslant \frac{1}{\varepsilon^2} \frac{1}{n^2} \sum_1^n \sigma_i^2 \to 0$$

である．このことから (2c. 3. 3) が成り立つ．もし，$\mu_i = \mu$ で $\sigma_i^2 = \sigma^2$ であるならば，定理の条件は自動的にみたされ，

$$\bar{X}_n \xrightarrow{P} \mu$$

を得る．

（ii）**ヒンチンの定理（大数の弱法則）**　$\{X_i\}$，$i = 1, 2, \cdots$ が互いに独立で同一の分布に従い，$E(X_i)$ が存在するとしよう．このとき

$$E(X_i) = \mu < \infty \Rightarrow \bar{X}_n \xrightarrow{P} \mu \tag{2c. 3. 4}$$

が成り立つ．

チェビシェフの定理のときのように2次の積率の存在は仮定されないが，観測値は互いに独立で同一分布に従うと考えられていることに注意するのは興味あることである．この定理の証明で特性関数の収束に基づくものは，この章の終りの問題 **4. 3** で与えられる．この定理をより強くして測度論の難解な命題に基づくものは (iv) で与えられる．しかし，ここではマルコフによる（初等的な結果だけを用いる）簡単な証明をしておく．

$i = 1, 2, \cdots, n$ と一定の δ にたいして一対の新しい確率変数を定義する：

$$Y_i = X_i,\quad Z_i = 0,\qquad |X_i| < \delta n \text{ のとき},$$
$$Y_i = 0,\quad Z_i = X_i,\qquad |X_i| \geqslant \delta n \text{ のとき}.$$

したがって，$X_i = Y_i + Z_i$ である．$i = 1, 2, \cdots, n$ にたいして $E(Y_i) = \mu_n$ としよう．$E(X_i) = \mu$ であるから，任意の ε にたいしてnを十分大きくとれば，

$$|\mu_n - \mu| < \varepsilon \tag{2c. 3. 5}$$

が成り立つ．いま，

$$V(Y_i) = \int_{-\delta n}^{\delta n} x^2 dF(x) - \mu_n^2 \leqslant \delta n \int_{-\delta n}^{\delta n} |x| dF(x) \leqslant b\delta n$$

が成立する，ただし $b = E|X|$ は仮説により存在する．チェビシェフの不等式を用いると

$$P\left(\left|\frac{1}{n}\sum_1^n Y_i - \mu_n\right| \geqslant \varepsilon\right) \leqslant \frac{b\delta}{\varepsilon^2}$$

を得，また (2c. 3. 5) を用いると

$$P\left(\left|\frac{1}{n}\sum_1^n Y_i - \mu\right| \geqslant 2\varepsilon\right) \leqslant \frac{b\delta}{\varepsilon^2}$$

を得る．ところで n を十分大きくとると，

$$P(Z_i \neq 0) = \int_{|x|\geqslant \delta n} dF(x) \leqslant \frac{1}{\delta n}\int_{|x|\geqslant \delta n} |x|\, dF(x) \leqslant \frac{\delta}{n}$$

となるから，

$$P\left(\sum_1^n Z_i \neq 0\right) \leqslant \sum_1^n P(Z_i \neq 0) \leqslant \delta \qquad \text{(2c. 3. 6)}$$

を得る．そこで，

$$\begin{aligned} P\left(\left|\frac{1}{n}\sum X_i - \mu\right| \geqslant 4\varepsilon\right) &= P\left[\left|\left(\frac{1}{n}\sum Y_i - \mu\right) + \frac{1}{n}\sum Z_i\right| \geqslant 4\varepsilon\right] \\ &\leqslant P\left[\left|\frac{1}{n}\sum Y_i - \mu\right| \geqslant 2\varepsilon\right] + P\left[\left|\frac{1}{n}\sum Z_i\right| \geqslant 2\varepsilon\right] \\ &\leqslant \frac{b\delta}{\varepsilon^2} + P(\sum Z_i \neq 0) \leqslant \frac{b\delta}{\varepsilon^2} + \delta \qquad \text{(2c. 3. 7)} \end{aligned}$$

を考えよう．ここで δ は任意であるから最後の値は小さくでき，求める結果が証明される．

(2c. 3. 6) で不等式

$$P(\sum Z_i \neq 0) \leqslant \sum P(Z_i \neq 0)$$

を利用したこと，また，(2c. 3. 7) で同じような不等式を利用したことを注意しておくがこれは簡単に証明できよう．

(iii) **コルモゴロフの定理 1 (大数の強法則)**　$\{X_i\}$，$i = 1, 2, \cdots$ を互いに独立な確率変数の系列とし $E(X_i) = \mu_i$，$V(X_i) = \sigma_i^2$ とする．このとき

$$\sum_1^\infty \frac{\sigma_i^2}{i^2} < \infty \Rightarrow \bar{X}_n - \bar{\mu}_n \xrightarrow{\text{a.s.}} 0 \qquad \text{(2c. 3. 8)}$$

が成り立つ，すなわち，系列 $X_1, X_2, \cdots$ は大数の強法則に従う．

これを証明するために確率変数 $Y_i = X_i - \mu_i$ を考え，ハイエク＝レニュイの不等式を適用しよう．ハイエク＝レニュイの不等式は，この章の終りの問題 3. 3 で証明する．この不等式は

$$P\left\{\max_{m\leqslant i\leqslant n} c_i|Y_1 + \cdots + Y_i| \geqslant \varepsilon\right\} \leqslant \frac{1}{\varepsilon^2}\left(c_m^2 \sum_1^m \sigma_i^2 + \sum_{m+1}^n c_i^2\sigma_i^2\right)$$

というものである．$c_i = 1/i$ とすることによって

$$P\left\{\max_{m\leqslant i\leqslant n} |\bar{Y}_i| \geqslant \varepsilon\right\} \leqslant \frac{1}{\varepsilon^2}\left(\frac{1}{m^2}\sum_1^m \sigma_i^2 + \sum_{m+1}^n \frac{\sigma_i^2}{i^2}\right)$$

を得る．$n \to \infty$ とすると，

$$P\left\{\max_{m\leqslant i} |\bar{Y}_i| \geqslant \varepsilon\right\} \leqslant \frac{1}{\varepsilon^2}\left(\frac{1}{m^2}\sum_1^m \sigma_i^2 + \sum_{m+1}^\infty \frac{\sigma_i^2}{i^2}\right) \qquad \text{(2c. 3. 9)}$$

となる．$\sum \sigma_i^2/i^2$ は収束するから，(2c. 3. 9) から

$$\lim_{m\to\infty} P\left\{\max_{m\leqslant i} |\bar{Y}_i| \geqslant \varepsilon\right\} = 0$$

が導かれる．しかしこれは

$$P\left\{\lim_{m\to\infty} \bar{Y}_m = 0\right\} = 1$$

ということであり，求める結果は証明された．

この定理で確率変数の独立性を，平均零の変数の直交性，すなわち，すべての $i \neq j$ にたいして $\mathrm{cov}(X_i, X_j) = 0$ におきかえるなら，このときは（より強い条件を必要とし）

$$\sum_1^\infty \frac{\sigma_i^2(\log i)^2}{i^2} < \infty \Rightarrow \bar{X}_n - \bar{\mu}_n \xrightarrow{\text{a.s.}} 0. \qquad \text{(2c. 3. 10)}$$

直交変数に関するこれおよび関連する結果の証明については Doob (1953) を参照されたい．

(iv) **コルモゴロフの定理 2 (大数の強法則)**　$X_1, X_2, \cdots$ を互いに独立で同一の分布に従う変数の列とする．$\bar{X}_n \xrightarrow{\text{a.s.}} \mu$ であるための必要十分条件は，$E(X_i)$ が存在しそれが μ に等しいことである．

必要性：E_n を $|X_n| \geqslant n$ である事象とする．このとき $\bar{X}_n \xrightarrow{\text{a.s.}} \mu$ であるから

$$\frac{X_n}{n} = \bar{X}_n - \frac{n-1}{n}\bar{X}_{n-1} \xrightarrow{\text{a.s.}} 0 \qquad \text{(2c. 3. 11)}$$

が成り立つ．(2c. 3. 11) から無限に多くの事象 E_n が起こる確率が零であることが導かれる．さらに，X_n の独立性から E_n の独立性が導かれ，ボレル＝カンテリの補助定理（付録 2B で証明する）によって

$$\sum_{n=1}^\infty P(|X_n| \geqslant n) = \sum_{n=1}^\infty P(E_n) < \infty \qquad \text{(2c. 3. 12)}$$

であることがわかる．いま，$p_j = P(|X| \geqslant j)$ とする．このとき

$$\begin{aligned} E(|X|) &\leqslant 1(1 - p_1) + 2(p_1 - p_2) + \cdots \\ &= 1 + p_1 + p_2 + \cdots = 1 + \sum_{n=1}^\infty P(E_n) \qquad \text{(2c. 3. 13)} \end{aligned}$$

である．(2c. 3. 12) から (2c. 3. 13) の最後の式は有限である．したがって，$E(|X|)$ が存在し，その十分条件から $E(X)=\mu$ が導ける．

十分性：先端が切れた変数の系列，すなわち

$$X_n^* = \begin{cases} X_n & |X_n| < n \text{ について} \\ 0 & |X_n| \geqslant n \text{ について} \end{cases}$$

を考える．このとき，

$$V(X_n^*) \leqslant E(X_n^{*2}) = \int_{-n}^{+n} x^2 dF(x) \leqslant \sum_{k=1}^{n-1} (k+1)^2 P(k \leqslant |X| < k+1)$$

および

$$\sum_{n=1}^{\infty} \frac{V(X_n^*)}{n^2} \leqslant \sum_{n=1}^{\infty}\sum_{k=0}^{n-1} \frac{(k+1)^2}{n^2} P(k \leqslant |X| < k+1)$$

$$\leqslant \sum_{=1}^{\infty} P(k-1 \leqslant |X| < k) k^2 \sum_{n=k}^{\infty} \frac{1}{n^2}$$

$$\leqslant 2\sum_{=1}^{\infty} kP(k-1 \leqslant |X| < k) \qquad \left(\sum_{n=k}^{\infty} \frac{1}{n^2} < \frac{1}{k^2} + \frac{1}{k} < \frac{2}{k} \text{ であるから}\right)$$

$$\leqslant 2[1 + E(|X|)] < \infty \qquad (E(X) \text{ が存在することから})$$

が得られる．したがって，(iii) で証明されたコルモゴロフの定理により，系列 X_n^* は大数の法則に従う，すなわち，

$$\bar{X}_n^* - \frac{1}{n}\sum_{1}^{n} E(X_i^*) \xrightarrow{\text{a.s.}} 0 \tag{2c. 3. 14}$$

である．次に，$n \to \infty$ のとき $E(X_n^*) \to E(X_n) = \mu$ であるから

$$\frac{1}{n}\sum_{1}^{n} E(X_i^*) \to \mu, \qquad n \to \infty \text{ のとき} \tag{2c. 3. 15}$$

が成り立つ．(2c. 3. 14) と (2c. 3. 15) を組み合わせることによって

$$\bar{X}_n^* \xrightarrow{\text{a.s.}} \mu$$

を得る．

次に，X_n と X_n^* が同等な系列であること，すなわち，

$$P\{X_n \neq X_n^*,\ n \geqslant N \text{ のある } n \text{ にたいして}\} \to 0, \qquad N \to \infty \text{ のとき}$$

であることを証明しよう．これから X_n^* が大数の強法則に従うなら X_n も強法則に従うこと，かつ，その極限が同じであることが導かれる．

$$P\{X_n \neq X_n^*,\ \text{ある } n \geqslant N \text{ にたいして}\}$$

$$\leqslant \sum_{n \geqslant N} P\{X_n \neq X_n^*\} = \sum_{n=N}^{\infty} P\{|X_n| \geqslant n\}$$

$$\leqslant \sum_{n=N}^{\infty} (n-N+1) P\{n \leqslant |X_n| < (n+1)\}$$

$$\leqslant \sum_{n=N}^{\infty} nP\{n \leqslant |X| < (n+1)\} \quad (\text{すべての } X_n \text{ が同じ分布関数をもつことから})$$

$$\leqslant \int_{|x| \geqslant N} |x| dF(x) \to 0 \quad N \to \infty \text{ のとき}$$

を考えればよい．

2c. 4　分布関数列の収束

$\{F_{X_n}\}$ または簡単に $\{F_n\}$，$n=1,2,\cdots$ によって，確率変数 $\{X_n\}$，$n=1,2,\cdots$ の分布関数の列をあらわそう．

定義　確率変数の列 $\{X_n\}$ は，F のすべての連続点において $n \to \infty$ のとき $F_n \to F$ であるならば，F_X（単に F でもあらわす）を分布関数とする確率変数へ，分布として（法則）収束するという．このような収束を $X_n \xrightarrow{L} X$ であらわす．

読者は実直線 $(-\infty, \infty)$ 上のいたるところで密であるような点の集合 S にたいして，$F_n \to F$ であることをいえば十分であると考えるであろう．

その近づいて行く先の分布 F を X_n の極限分布または漸近分布という．統計の応用において，極限分布は重要な役割を演じる．大きさ n の標本から計算され，その実際の分布を見出すのはむずかしい統計量を確率変数 X_n であらわす．そのような場合には少なくとも n の値が大きければ，その極限分布によって近似することができる．次の結果は極限分布の研究において重要なものである．

（i）**ヘリーの補助定理**　分布関数の系列はいずれも弱コンパクトである．すなわち，ある関数（分布関数である必要はない）にそのすべての連続点において収束する部分列が存在する．

$D=\{r_k\}$ をすべての有理数の集合とする．$F_n(r_1)$ が有界であることから，ボルツァノ＝ワイエルシュトラスのコンパクト定理によって収束部分列が存在する．列 $\{F_{n_1}(x)\}$ が特定の値 $x=r_1$ で収束すると考えよう．列 $\{F_{n_1}(x)\}$ から他の部分列 $\{F_{n_2}(x)\}$ を同様な方法でとり出すことができる．ここで $\{F_{n_2}(x)\}$ は $x=r_2$ で収束し，もちろん $x=r_1$ ででも収束するであろう．以下同様である．そこで $\{F_{n_1}\}$ の第1の要素，$\{F_{n_2}\}$ の第2の要素，… で与えられる列を考えよう．そのような関数の列 $\{F_s\}$ はすべての $x \in D$ にたいして必ず収束し，極限関数 F_D はすべての $x \in D$ で定義され，有界で非減少である．任意の x にたいして

$$F(x) = \underset{r_s<x}{\text{上限}}\ F_D(r_i) \tag{2c. 4. 1}$$

とする．定義よりFは左から連続で有界で非減少である．

定められた部分列は (2c. 4. 1) で定義した $F(x)$ にFのすべての連続点で実際に収束することを示そう．xがそのような点であるとする．このとき $x_i' < x < x_i''$ であり，$i \to \infty$ のとき $F(x_i'') - F(x_i') \to 0$ であるような有理数 (x_i', x_i'') の列をみつけることができる．このような任意の対 (x_i', x_i'') にたいし各sについて自明な関係

$$F_s(x_i') \leqslant F_s(x) \leqslant F_s(x_i'') \tag{2c. 4. 2}$$

を得る．ここで $\{F_s(x)\}$ は $F_D(x)$ に収束する列である．(2c. 4. 2) の関係の極限をとることによって，各iにたいして

$$F_D(x_i') \leqslant \underline{\lim}\, F_s(x) \leqslant \overline{\lim}\, F_s(x) \leqslant F_D(x_i'')$$

を得る．$F_D(x_i'') - F_D(x_i')$ はどのようにも小さくできることから $\lim F_s(x)$ が存在し，(2c. 4. 1) で定義した $F(x)$ に等しい．$F(x)$ もまた区間 $[F_D(x_i'), F_D(x_i'')]$ にあることがわかる．

（ii）ヘリー＝ブレイの定理　<u>すべての有界連続関数gについて</u>

$$F_n \to F \Rightarrow \int g dF_n \to \int g dF.$$

Fの2つの連続点 a, b $(a < b)$ をとれば

$$\int_{-\infty}^{\infty} g dF_n - \int_{-\infty}^{\infty} g dF$$
$$= \int_{-\infty}^{a} g(dF_n - dF) + \int_a^b g(dF_n - dF) + \int_b^{\infty} g(dF_n - dF) \tag{2c. 4. 3}$$

と書ける．

$|g| < c$ とする．このことから (2c. 4. 3) の右辺第1項の積分の絶対値は

$$\left|\int_{-\infty}^{a} g dF_n - \int_{-\infty}^{a} g dF\right| < c\int_{-\infty}^{a} dF_n + c\int_{-\infty}^{a} dF = c[F_n(a) + F(a)]$$

となる．もしaが十分小さいとすると $F(a)$ は小さくまたすべての $n > n_0$ にたいして $F_n(a)$ も小さい．このことから，適当に選んだaと n_0 にたいして $c[F_n(a) + F(a)] < \varepsilon/5$ となる．同様に (2c. 4. 3) の右辺3番目の積分の絶対値は適当に選んだbと n_0 にたいして $< \varepsilon/5$ である．

有限区間 (a, b) で，gは一様連続である．(a, b) をn個の区間に分け

$$x_0 = a < x_1 < \cdots < x_{m-1} < b = x_m$$

とする．ここで $x_1, \cdots, x_{m-1}$ はFの連続点で，すべてのiにたいして区間 $(x_i < x < x_{i+1})$ で一様に $[g(x) - g(x_i)] < \varepsilon/5$ である．関数

$$g_m(x) = g(x_i), \qquad x_i \leqslant x < x_{i+1}$$

を定義する．このとき $n \to \infty$ で

$$\int_a^b g_m(x) dF_n = \sum g(x_i)[F_n(x_{i+1}) - F_n(x_i)]$$
$$\to \sum g(x_i)[F(x_{i+1}) - F(x_i)] = \int_a^b g_m dF$$

であるから，任意のmにたいして

$$\left|\int_a^b g_m(dF_n - dF)\right| < \varepsilon/5$$

が十分大きいnにたいして成り立つ．しかし，

$$\left|\int_a^b g dF_n - \int_a^b g dF\right| = \left|\int_a^b (g - g_m) dF_n + \int_a^b g_m(dF_n - dF) + \int_a^b (g - g_m) dF\right|$$
$$< \int_a^b \frac{\varepsilon}{5} dF_n + \frac{\varepsilon}{5} + \int_a^b \frac{\varepsilon}{5} dF < \frac{3}{5}\varepsilon$$

が十分大きいnにたいして成り立つ．このことから差 (2c. 4. 3) の絶対値は $< \varepsilon$ であり，求める結果が得られた．

（iii）拡張したヘリー＝ブレイの定理　<u>$F_n \to G$ とする．ここでGは必ずしも (ii) で仮定したような分布関数である必要はない．また，gは $g(\pm\infty) = 0$ であるような連続関数とする．$\Rightarrow \int g dF_n \to \int g dG.$</u>

式 (2c. 4. 3) を考える．そして右辺の第2項が，Gが分布関数であると否とにかかわらず，(ii) で使ったのと同じ論法で，零に近づくことに注意する．gが有界であるから $\int g dF_n$ と $\int g dG$ が存在する．aを十分小さく選ぶことによって第1項は $(-\infty, a)$ の範囲にあるxにたいして，nの値に関係なく $2\sup|g(x)|$ でおさえられ，この限界は，$g(-\infty) = 0$ であり，かつgが連続であることから小さい値になる．同様に，bを十分大きく選ぶことによって，第3項を小さくすることができる．以上から差 (2c. 4. 3) は十分大きなnをとることによって，どれだけでも小さくすることができる．

（iv）連続性の定理　<u>$f_n(t)$ を X_n の特性関数とする．もし $X_n \xrightarrow{L} X$ であるならば $f_n(t) \to f(t)$ である．ここで $f(t)$ は X の特性関数である．もし $f_n(t) \to f(t)$ で極限関数 $f(t)$ が $t = 0$ で連続であるならば，$X_n \xrightarrow{L} X$ で $f(t)$ は X の特性関数である．</u>

最初の命題はヘリー＝ブレイの定理から直接得られる．第2の命題を証明するために $t = 0$ での $f(t)$ の連続性を用いる．

F_nをX_nの分布関数とし，部分列 $\{F_n\}$ は有界非減少左連続関数Gに (i) のヘリーの補助定理の場合のように収束するものとする．このとき

$$\int_0^v f_n(t) dt = \int_0^v \int e^{itx} dF_m dt$$
$$= \int dF_m \int_0^v e^{itx} dt = \int \frac{e^{ivx} - 1}{ix} dF_m \tag{2c. 4. 4}$$

である．関数 $(e^{ivx}-1)/ix$ は $x\to\mp\infty$ のとき0に近づく．そこで (iii) の拡張したヘリー＝ブレイの定理を使うと，(2c. 4. 4) の両辺の極限をとることによって

$$\int_0^v f(t)dt=\int\frac{e^{ivx}-1}{ix}dG \qquad (2c.\ 4.\ 5)$$

を得る．両辺を v で割り $v\to 0$ という極限をとることによって，[$f(t)$ の $t=0$ における連続性から]

$$f(0)=\int\lim_{v\to 0}\frac{e^{ivx}-1}{ivx}dG=\int dG=G(\infty)-G(-\infty) \qquad (2c.\ 4.\ 6)$$

を得る．しかし $\lim f_m(0)=1=f(0)$ であるから G は分布関数である．$F_n\to G$ であることから (ii) のヘリー＝ブレイの定理によって

$$f(t)=\lim_{m\to\infty}f_m(t)=\lim_{m\to\infty}\int e^{itx}dF_m=\int e^{itx}dG$$

を得る．したがって $f(t)$ は特性関数であり，F_n のすべての収束する部分列にたいして G はもちろん一意的である．よって $F_n\to G$ である．

(v) (ii) と (iv) を結合することによって，次の3つは同値である：

(a) $F_n\to F$,　(b) $f_n(t)\to f(t)$,　(c) $\int gdF_n\to\int gdF$ 　　(2c. 4. 7)

ここで g と $f(t)$ は (ii) と (iv) でそれぞれ定義したものである．

(v) から，もし $X_n\xrightarrow{L}X$ ならば，a と b とを定数とするとき $aX_n+b\xrightarrow{L}aX+b$ であることが容易に示せる．さらに $X_n\xrightarrow{P}c$ (定数) $\Leftrightarrow F_n\to F_c$ である．ここで F_c は確率1で値 c をとるような確率変数 X の分布関数で，すなわち，階段関数であり，$x\leqslant c$ のとき $F_c(x)=0$，$x>c$ のとき $F_c(x)=1$ である．

(vi) ポリヤの定理　もし $F_n\to F$ で F が連続であるならばその収束は一様である．すなわち

$$\lim_{n\to\infty}\sup_x|F_n(x)-F(x)|=0. \qquad (2c.\ 4.\ 8)$$

証明は容易である．

(vii) g が連続関数で $F_n\to F$ とする．このとき次の命題が成り立つ．

(a) $\liminf_{n\to\infty}\int|g|dF_n\geqslant\int|g|dF$,

(b) $\int|g|dF_n$ は一様積分可能 $\Rightarrow\int gdF_n\to\int gdF$.

これを示すには

$$\int|g|dF_n\geqslant\int_{-c}^{c}|g|dF_n$$

$$\liminf_{n\to\infty}\int|g|dF_n\geqslant\lim\int_{-c}^{c}|g|dF_n$$

$$=\int_{-c}^{c}|g|dF \qquad \text{任意の } c \text{ にたいして,}$$

であることから (a) が示される．一様に積分可能であるから n と独立に ε が存在し

$$\varepsilon+\int_c^d|g|dF_n\geqslant\int|g|dF_n\geqslant\int_{c'}^{d'}|g|dF_n,\qquad c'<c,\ d'>d$$

である．

極限をとると，c, d, c', d' が有限であることから

$$\varepsilon+\int_c^d|g|dF\geqslant\int_{c'}^{d'}|g|dF$$

であることがわかる．ここで c' と d' は任意の値である．このことから $|g|$ したがって g は dF にたいして積分可能である．そこで

$$\int gdF_n-\int gdF=\int_{-c}^{c}g(dF_n-dF)+\int_{|x|>c}gdF_n-\int_{|x|>c}gdF \qquad (2c.\ 4.\ 9)$$

である．

(2c. 4. 9) の右辺のすべての項は c と n とを十分大きくとることによって小さい値にすることができる．このことから (b) が証明された．

(vii-A) γ_{rn} を X_n の r 次の絶対積率とする．もし，ある $r>0$ があって，すべての大きい n について γ_{rn} がある定数より大きくないならば，次のことがいえる．

(a) 分布関数 G に収束する部分列 F_m が存在する．

(b) α_{sm} と α_s を F_m と G の s 次積率とする (s と r とはともに整数でなくともよい) とき，$s<r$ にたいして $\alpha_{sm}\to\alpha_s$ である．

ヘリーの補助定理により，$F_m\to G$ で G は非減少有界関数である．もし c [(>0)] を G の連続点とすると，十分大きい m にたいして

$$\int_{|x|>c}dF_m\leqslant c^{-r}\int_{|x|>c}|x|^r dF_m\leqslant c^{-r}k \qquad (2c.\ 4.\ 10)$$

である．(2c. 4. 10) の第1項は

$$1-[F_m(c)-F_m(-c)]\leqslant c^{-r}k$$

である．$m\to\infty$ として極限をとると

$$1-[G(c)-G(-c)]\leqslant c^{-r}k$$

を得る．$c\to\infty$ とすると

$$1-[G(\infty)-G(-\infty)]\leqslant 0$$

または $G(\infty)-G(-\infty)\geqslant 1$ である．これは G が分布関数であることを示しており，(a) が証明された．

次に

$$\int_{|x|>c}|x|^s dF_m \leqslant c^{s-r}\int_{|x|>c}|x|^r dF_m < c^{s-r}k$$

であることから $|x|^s$ はもし $s<r$ であるならば一様積分可能である．したがって，(b) は (vii) を使うことによって証明される．

(viii-B) α_{sn} を F_n の s 次積率で $\alpha_{sn}\to\alpha_s$ とする．ここで α_s と α_{sn} とはすべての s とすべての大きい n にたいして有限であるとする．このとき次の関係が得られる：

(a) $F_n\to F \Rightarrow \alpha_s$ は F の s 次積率

(b) α_s が一意的な分布関数 F を決定する $\Rightarrow F_n\to F$.　(2c. 4. 11)

結果 (a) は (vii) から直接導かれる．さらに (vii) によって分布関数 G に収束する部分列が存在し，α_s は G の積率である．もし α_s が一意的な分布関数を決定するとすると，すべての部分列は同一の極限値をもつ．このことから $F_n\to F$ である．

積率が与えられたとき F が一意的であるための十分条件は

$$\sum_1^\infty(\alpha_{2s})^{-1/2s}=\infty$$

であるから，(b) は

$$\sum_1^\infty(\alpha_{2s})^{-1/2s}=\infty \Rightarrow F_n\to F$$

とも書ける．

(ix) $\{X_n, Y_n\}$，$n=1,2,\cdots$ を変数の対の列とすると

$$|X_n-Y_n|\xrightarrow{P}0,\quad Y_n\xrightarrow{L}Y \Rightarrow X_n\xrightarrow{L}Y \qquad (2c.\ 4.\ 12)$$

である．すなわち，X_n の極限分布は存在しそれは Y の極限分布と同一のものである．

X_n の分布関数を F_{X_n} とあらわし，Y の分布関数を F_Y とあらわす．$Y_n-X_n=Z_n$ とし x を F_Y の連続点としよう．このとき

$$\begin{aligned}F_{X_n}(x)&=P(X_n<x)=P(Y_n<x+Z_n)\\&=P(Y_n<x+Z_n,\ Z_n<\varepsilon)+P(Y_n<x+Z_n,\ Z_n\geqslant\varepsilon)\\&\leqslant P(Y_n<x+\varepsilon)+P(Z_n\geqslant\varepsilon)\end{aligned}$$

である．極限をとると

$$\limsup_{n\to\infty}F_{X_n}\leqslant F_Y(x+\varepsilon)$$

を得る．同様に

$$\liminf_{n\to\infty}F_{X_n}\geqslant F_Y(x-\varepsilon)$$

である．ε が任意の値であり，x が F_Y の連続点であることから $\varepsilon\to0$ にすることによって

$$\lim_{n\to\infty}F_{X_n}(x)=F_Y(x)$$

で (2c. 4. 12) が証明された．特別な場合として

$$X_n\xrightarrow{P}X \Rightarrow X_n\xrightarrow{L}X$$

を得る．

(x) $\{X_n, Y_n\}$，$n=1,2,\cdots$ を確率変数の対の列とする．このとき次の関係が得られる：

(a) $X_n\xrightarrow{L}X$，$Y_n\xrightarrow{P}0 \Rightarrow X_nY_n\xrightarrow{P}0$.

(b) $X_n\xrightarrow{L}X$，$Y_n\xrightarrow{P}c \Rightarrow X_n+Y_n\xrightarrow{L}X+c$

$\Rightarrow X_nY_n\xrightarrow{L}cX$

$\Rightarrow X_n/Y_n\xrightarrow{L}X/c$，　もし $c\neq0$ ならば，

(c) $X_n\xrightarrow{L}X$，$Y_n\xrightarrow{P}c \Rightarrow (X_n, Y_n)$ の同時分布の極限が存在しそれは (X, c) の同時分布の極限に等しい．

(d) $X_n-Y_n\xrightarrow{P}0$，$X_n\xrightarrow{L}X \Rightarrow Y_n\xrightarrow{L}X$.

(a) を示すために，次のことを考える：

$$\begin{aligned}P(|X_nY_n|>\varepsilon)&=P(|X_nY_n|>\varepsilon, |Y_n|\leqslant\varepsilon/k)+P(|X_nY_n|>\varepsilon, |Y_n|>\varepsilon/k)\\&\leqslant P(|X_n|>k)+P(|Y_n|>\varepsilon/k)\\&\Rightarrow \limsup_{n\to\infty}P(|X_nY_n|>\varepsilon)\leqslant P(|X|>k)\end{aligned}$$

ある固定した k にたいして，

(2c. 4. 13)

ここで k は任意である．したがって k を大きく選ぶことによって，(2c. 4. 13) の右辺を小さくすることができる．このことから $P(|X_nY_n|>\varepsilon)\to0$ である．

(b) を示すために，$X_n\xrightarrow{L}X \Rightarrow X_n+c\xrightarrow{L}X+c$ とする．このとき

$$(X_n+Y_n)-(X_n+c)=Y_n-c\xrightarrow{P}0.$$

このことから (ix) より，$X_n+Y_n\xrightarrow{L}X+c$ である．さらに $X_n\xrightarrow{L}X \Rightarrow cX_n\xrightarrow{L}cX$ である．次に，(a) と (x) より

$$X_nY_n-cX_n=X_n(Y_n-c)\xrightarrow{P}0.$$

このことから (ix) より，$X_nY_n\xrightarrow{L}cX$．同様に $X_n/Y_n\xrightarrow{L}X/c$ である．

(c) を示すために X の分布の連続点 x と $y>c$ とを考える．このとき

$$P(X_n < x) \geqslant P(X_n < x,\ Y_n \leqslant y) = P(X_n < x) - P(X_n < x,\ Y_n > y)$$
$$\geqslant P(X_n < x) - P(Y_n > y)$$

である．極限をとると，

$$\lim_{n\to\infty} P(X_n < x, Y_n < y) = P(X < x), \qquad \text{もし } y > c \text{ ならば}$$

である．同様に

$$\lim_{n\to\infty} P(X_n < x,\ Y_n < y) = 0, \qquad \text{もし } y < c \text{ ならば}$$

である．$y = c$ のとき極限は不定で，任意の x にたいして (x, c) は (X, c) の同時分布の不連続点であることに注意しておく．

(d) は (ix) と同値である．

(xi) $\{X_n^{(1)}, \cdots, X_n^{(k)}\}$，$n = 1, 2, \cdots$ を確率変数のベクトルの列とする．任意の実数 $\lambda_1, \cdots, \lambda_k$ にたいして

$$\lambda_1 X_n^{(1)} + \cdots + \lambda_k X_n^{(k)} \xrightarrow{L} \lambda_1 X^{(1)} + \cdots + \lambda_k X^{(k)}$$

とする．ここで $X^{(1)}, \cdots, X^{(k)}$ の同時分布を $F(x_1, \cdots, x_k)$ とする．このとき，$X_n^{(1)}, \cdots, X_n^{(k)}$ の同時分布関数の極限が存在し，それは $F(x_1, \cdots, x_k)$ に等しい．

この証明は明らかである．なぜなら

$$\lambda_1 X^{(1)} + \cdots + \lambda_n X_n^{(k)} \xrightarrow{L} \lambda_1 X^{(1)} + \cdots + \lambda_k X^{(k)}$$

は対応する特性関数が収束することを意味している．もし $f_n(t_1, \cdots, t_k)$ が $(X_n^{(1)}, \cdots, X_n^{(k)})$ の特性関数で $f(t_1, \cdots, t_k)$ が $(X^{(1)}, \cdots, X^{(k)})$ の特性関数であるとすると

$$f_n(t_1\lambda_1, \cdots, t_k\lambda_k) \to f(t_1\lambda_1, \cdots, t_k\lambda_k)$$

である．$\lambda_1, \cdots, \lambda_k$ が任意であることから

$$f_n(t_1, \cdots, t_k) \to f(t_1, \cdots, t_k)$$

で，これは求める結果を示す．

(xii) もし g が連続関数で $X_n \xrightarrow{L} X$ であるとすると $g(X_n) \xrightarrow{L} g(X)$ である．

$\cos tg(x)$ と $\sin tg(x)$ が x の有界連続関数であることから (ii) によって次式を得る．

$$\begin{aligned} E[e^{itg(X_n)}] &= \int e^{itg(x)} dF_n(x) \\ &= \int \cos tg(x) dF_n(x) + i\int \sin tg(x) dF_n(x) \\ &\to \int \cos tg(x) dF(x) + i\int \sin tg(x) dF(x), \\ &= E[e^{itg(x)}] \end{aligned}$$

したがって，(2c. 4. 7) によって証明される．

(xiii) g が連続関数で $X_n \xrightarrow{P} X$ であるならば

$$g(X_n) \xrightarrow{P} g(X) \tag{2c. 4. 14}$$

である．

I を $P(X \in I) = 1 - \eta/2$ であるような有限区間とし，$P(|X_n - X| < \delta) > 1 - \eta/2$ であるような $n \geqslant n_0$ をとる．すると，任意の $X \in I$ にたいして $|X_n - X| < \delta$ であるとき $|g(X_n) - g(X)| < \varepsilon$ である．このことから次の関係がわかる．

$$\begin{aligned} P(|g(X_n) - g(X)| < \varepsilon) &\geqslant P(|X_n - X| < \delta,\ X \in I) \\ &\geqslant P(|X_n - X| < \delta) - P(X \notin I) \\ &\geqslant 1 - \eta \qquad n \geqslant n_0 \text{ にたいして．} \end{aligned}$$

(2c. 4. 14) の重要な応用として次のことがある．すなわち，$X_n \xrightarrow{P} c$ (定数) のとき，g が連続関数であるという条件で $g(X_n) \xrightarrow{P} g(c)$ である．

(xiv) g を連続関数とすると

$$X_n - Y_n \xrightarrow{P} 0, \quad Y_n \xrightarrow{L} Y \Rightarrow g(X_n) - g(Y_n) \xrightarrow{P} 0$$

である．

注意　Mann and Wald (1943) は，より一般的な条件のもとで (xii)—(xiv) を証明した．その条件は g をボレル可測と仮定するが，g の不連続の集合に制限を加えるというものである．

また，(i) から (xiv) は X_n, Y_n がベクトル変数であるときにも成り立つことを注意しておこう．証明はまったく同様である．

(xv) 密度関数に関する収束定理 (Scheffe (1947))　確率変数 X_n の確率密度関数を $p_n(x)$ とし，$n \to \infty$ のとき $p_n(x) \to p(x)$ とする．このとき次の関係が得られる．

(a) $p(x)$ が密度関数 $\Rightarrow \int_R |p_n(x) - p(x)| dx \to 0.$

(b) $|p_n(x)| < q(x)$ であり $\int q(x) dx$ が存在する $\Rightarrow p(x)$ は密度関数である．また (a) により $\int |p_n(x) - p(x)| dx \to 0$ である．

(a) と (b) のいずれも，すべてのボレル集合体 S にたいして一様に

$$\int_S [p_n(x) - p(x)] dx \to 0$$

である．したがって，$X_n \xrightarrow{L} X$ であることを意味している．ここで X は密度関数 $p(x)$ で定義される確率変数である．

いま，$\delta_n(x) = p_n(x) - p(x)$ とし

$$\delta_n^+ = \frac{1}{2}(\delta_n + |\delta_n|) \geqslant 0, \qquad \delta_n^- = \frac{1}{2}(\delta_n - |\delta_n|) \leqslant 0$$

をそれぞれ δ_n の正の部分，負の部分と定義する．明らかに R に属するほとんどすべての x にたいして $\delta_n \to 0$ のとき $\delta_n^+ \to 0$，$\delta_n^- \to 0$ である．

$p_n(x)$ と $p(x)$ とが密度関数で，$\delta_n = \delta_n^+ + \delta_n^-$ であることから

$$0 = \int_R [p_n(x) - p(x)]dx = \int_R \delta_n dx = \int_R \delta_n^+ dx + \int_R \delta_n^- dx \qquad (2c.4.15)$$

である．$\delta(x)$ は正でも負でも $p_n(x) - p(x) = \delta_n(x)$ が $\delta_n(x) \geqslant -p(x)$ である．このことから $\delta_n^-(x)$ は条件 $0 \geqslant \delta_n^-(x) \geqslant -p(x)$ を満足する．すなわち，$|\delta_n^-(x)| < p(x)$ で $\int p(x)dx$ が存在する．ルベーグ上限収束定理（付録 **2B**）を使うと，$\lim \delta_n^-$ が 0 であるから，

$$\int_R \delta_n^- dx \to \int_R \lim \delta_n^- dx = 0$$

が得られ，(2c.4.15) から

$$\int_R \delta_n^+ \to 0$$

また，このことから

$$\int_R |\delta_n| dx = \int_R \delta_n^+ dx - \int_R \delta_n^- dx \to 0$$

を得，(a) が示される．結果 (a) はまた付録 **2B** の補助定理 **2B.3** の直接の応用として証明される．結果 (b) は上限収束定理から導かれ

$$1 = \int_R p_n dx \to \int_R p(x)dx$$

である．

(xvi)　X_n が離散変数で $P(X_n = r) = p_n(r)$，$r = 0, 1, 2, \cdots$ であるとする．$Y_n = (X_n - n\mu)/\sqrt{n}$ を考える．ここで μ は変数である．もし $\sqrt{n}\, p_n(\sqrt{n}\, y + n\mu) \to p(y)$ で $p(y)$ が確率変数 Y の密度関数であるならば $Y_n \xrightarrow{L} Y$ である．

証明は (xv) で与えたシェッフェの定理に基づく．詳細な証明は読者にまかせよう．この結果の応用に関しては Okamoto (1960) を参照されたい．

(xvii)　**離散分布の収束**　$P_n(t)$，$0 \leqslant t \leqslant 1$ を確率母関数の列とすると任意の添字 n にたいして，$P_n(t) = \sum p_{x_n} t^x$ で定義された離散確率分布が存在する．さらに $P(t) = \sum p_x t^x$ とする．このことから

$$p_{x_n} \to p_x \text{（各 } x \text{ にたいして）} \Leftrightarrow P_n(t) \to P(t) \text{（各 } t \text{ にたいして）}$$

が成り立つ．項 p_x は $\sum p_x = 1$ であるときのみ分布と定義される．その場合

$$p_{x_n} \to p_x \Leftrightarrow \sum_{x=0}^{\infty} |p_{x_n} - p_x| \to 0$$

である．

この証明は (xv) の証明と同様である．

2c.5　中心極限定理

互いに独立で同一の分布に従う確率変数の列 $\{X_n\}$，$n = 1, 2, \cdots$ にたいして，**2c.3** は $\mu = E(x)$ が存在するという条件のもとで弱い意味または強い意味で $\bar{X}_n \to \mu$ であることを示した．しかし，このことは大標本で $\bar{X}_n$ の分布をどのように近似すればよいかということに関しては考え方を示してはいない．このことを調べるために確率変数の列

$$Y_n = \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \qquad (2c.5.1)$$

を考えよう．ここで σ はある定数である．いくつかの条件のもとで極限分布関数は

$$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx \qquad (2c.5.2)$$

でこれは (2a.5.8) で示した平均 0，分散 1 の正規分布関数である．平均 μ，分散 σ^2 の正規分布のより一般的な確率密度関数は

$$\frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2} \qquad (2c.5.3)$$

である．(2b.4.11) で求めたようにして (2c.5.2) の特性関数を求めれば

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{itx} e^{-x^2/2} dx = e^{-t^2/2} \qquad (2c.5.4)$$

である．(2c.5.1) が正規確率変数に収束するのを証明するために使われる手法は，Y_n の特性関数 $f_n(t)$ を計算して，

$$\lim_{n\to\infty} f_n(t) = e^{-t^2/2}$$

を示すことである．このことに (2c.4.7) を用いて，求める結果が得られる．

次に，より一般的な問題，すなわち確率変数

$$Y_n = \frac{\sqrt{n}(\bar{X}_n - \mu_n)}{\sigma_n}$$

の収束に関する問題を考える．ここで $\{\mu_n, \sigma_n\}$，$n = 1, 2, \cdots$ は 2 個ずつの数の対の列である．

(i)　**リンドバーグ＝レヴィの定理（中心極限定理）**　$X_1, X_2, \cdots$ を独立で同一の分布に従う確率変数の列とし，$E(X_n) = \mu$ で $V(X_n) = \sigma^2 \neq 0$ が存在するものとする．このとき Y_n の分布関数は $\Phi(x)$ に収束する．ここで Y_n は (2c.5.1) で定義され，Φ

は (2c. 5. 2) で定義されたものである．

$f(t)$ を $X_i-\mu$ の特性関数とする．1次および2次のモーメントが存在するから

$$f(t)=1-\frac{1}{2}\sigma^2t^2+o(t^2) \tag{2c. 5. 5}$$

である．$Y_n=\sum_1^n(X_i-\mu)/\sqrt{n}\sigma$ の特性関数は，(2c. 5. 5) を用いると，

$$f_n(t)=\left[f\left(\frac{t}{\sigma\sqrt{n}}\right)\right]^n=\left[1-\frac{t^2}{2n}+o\left(\frac{t^2}{n}\right)\right]^n.$$

右辺の対数をとると，

$$\log\left[1-\frac{t^2}{2n}+o\left(\frac{t^2}{n}\right)\right]^n=n\log\left[1-\frac{t^2}{2n}+o\left(\frac{t^2}{n}\right)\right]\to\frac{-t^2}{2}$$

すなわち，

$$f_n(t)\to e^{-t^2/2}\qquad n\to\infty \text{ のとき}$$

となる．

極限分布が連続であることから Y_n の分布関数は実際に一様収束し，より一般的な結果，すなわち

$$\lim_{n\to\infty}[F_{Y_n}(x_n)-\Phi(x_n)]\to 0 \tag{2c. 5. 6}$$

が得られる．ここで x_n はどのような数列であってもよい．

(2c. 5. 6) は $\bar{X}_n$ の分布関数が十分大きな n にたいして平均 μ 分散 σ^2/n をもつ正規分布関数で近似することができることを意味している．これを中心極限定理という．

（ii）**リャプノフの定理（中心極限定理）**　$\{X_n\}$，$n=1,2,\cdots$ を互いに独立な確率変数の列とする．各 n にたいして $E(X_n)=\mu_n$，$E(X_n-\mu_n)^2=\sigma_n^2\neq 0$，$E|X_n-\mu_n|^3=\beta_n$ が存在するとしよう．さらに

$$B_n=\left(\sum_1^n\beta_i\right)^{1/3},\qquad C_n=\left(\sum_1^n\sigma_i^2\right)^{1/2}$$

とする．このとき，もし $n\to\infty$ のとき $\lim(B_n/C_n)=0$ ならば

$$Y_n=\frac{\sum_1^n(X_i-\mu_i)}{C_n}$$

の分布関数は $\Phi(x)$ に収束する．

この定理の証明はリンドバーグ＝レヴィ（中心極限定理）の定理の証明と同じで，Y_n の特性関数を展開し，$n\to\infty$ のときの極限を考えればよい．詳細に関しては Gnedenko (1962) を参照されたい．

（iii）**リンドバーグ＝フェラーの定理**　$\{X_n\}$ を互いに独立な確率変数の列とし G_i を X_n の分布関数とする．さらに $E(X_n)=\mu_n$ と $V(X_n)=\sigma_n^2\neq 0$ が存在するとする．

$$Y_n=\frac{\sum_1^n(X_i-\mu_i)}{C_n},\qquad C_n=\left(\sum_1^n\sigma_i^2\right)^{1/2}$$

とすると

$$\lim_{n\to\infty}\max_{1\le i\le n}\frac{\sigma_i}{C_n}=0,\qquad F_{Y_n}\to\Phi$$

は，すべての $\varepsilon>0$ にたいして

$$\lim_{n\to\infty}\frac{1}{C_n^2}\sum_{i=1}^n\int_{|x-\mu_i|>\varepsilon C_n}(x-\mu_i)^2dG_i(x)=0$$

のとき，しかもそのときにかぎって成り立つ．

この定理の証明と中心極限定理の他の形式のものについては Gnedenko (1962) と Gnedenko and Kolmogorov (1954) を参照されたい．

（iv）**多変量中心極限定理** (Varadarajan (1958), Wald and Wolfowiz (1944))　F_n を k 次元確率変数 $(X_n^{(1)},\cdots,X_n^{(k)})$，$n=1,2,\cdots$ の同時分布関数と定義し，F_{λ_n} を線形関数 $\lambda_1X_n^{(1)}+\cdots+\lambda_kX_n^{(k)}$ の分布関数と定義する．F_n が k 変数分布関数 F に収束するための必要十分条件は F_{λ_n} が各ベクトル λ に関してある極限に収束することである．

証明は Varadarajan (1958) を参照せよ．本書で考察される応用に関しては［(xi), 2c. 4］で証明した弱い定理で十分である．すなわち，

F_n と F_{λ_n} を (iv) と同様に定義する．F を k 次元確率変数 $X^{(1)},\cdots,X^{(k)}$ の同時分布関数とする．もし各ベクトル λ にたいして，$F_{\lambda_n}\to F_\lambda$ で F_λ が $\lambda_1X^{(1)}+\cdots+\lambda_kX^{(k)}$ の分布関数であるならば $F_n\to F$ である．

応用として，互いに独立で同一の分布に従う k 次元変数

$$\mathbf{U}_n'=(U_{1n},\cdots,U_{kn}),\qquad n=1,2,\cdots$$

を考える．これらには第1次モーメント，第2次モーメントが存在しそれぞれ $E(U_n)=\boldsymbol{\mu}$，$D(U_n)=\boldsymbol{\Sigma}$ とする．確率変数の列を

$$\bar{\mathbf{U}}_n=(\bar{U}_{1n},\cdots,\bar{U}_{kn}),\quad n=1,2,\cdots,\quad \bar{U}_{in}=\frac{1}{n}\sum_1^nU_{ij}$$

と定義する．$\sqrt{n}(\bar{\mathbf{U}}_n-\boldsymbol{\mu})$ の漸近分布は $N_k(0,\boldsymbol{\Sigma})$，すなわち，$k$ 変数正規分布で，平均0，分散行列 $\boldsymbol{\Sigma}$ の分布であり密度関数は

$$N_k(\mathbf{u}|0,\boldsymbol{\Sigma})=(2\pi)^{-k/2}|\boldsymbol{\Sigma}|^{-1/2}\exp\left\{-\frac{1}{2}[(\mathbf{u}-\boldsymbol{\mu})'\boldsymbol{\Sigma}(\mathbf{u}-\boldsymbol{\mu})]\right\} \tag{2c. 5. 7}$$

である．

このことを示すには，ある一定のベクトル $\boldsymbol{\lambda}$ にたいして $\sqrt{n}\boldsymbol{\lambda}'(\bar{\mathbf{U}}_n-\boldsymbol{\mu})$ の極限分布が $N_1(0,\boldsymbol{\lambda}'\boldsymbol{\Sigma}\boldsymbol{\lambda})$ となることを証明すればよい．なぜなら $\mathbf{U}$ の分布が $N_k(0,\boldsymbol{\Sigma})$ であれば

$N_1(0, \boldsymbol{\lambda}'\boldsymbol{\Sigma}\boldsymbol{\lambda})$ が $\boldsymbol{\lambda}'\boldsymbol{\Sigma}$ の分布であるからである．この最後の結果の証明のためには $\mathbf{U}$ の密度関数が (2c. 5. 7) であるときの $\boldsymbol{\lambda}'\mathbf{U}$ の特性関数の計算をすればよい．いま，1次元確率変数の列

$$U_n = \boldsymbol{\lambda}'\mathbf{U}_n = \lambda_1 U_{1n} + \cdots + \lambda_k U_{kn}$$

を考える．ここで $E(U_n) = \boldsymbol{\lambda}'\boldsymbol{\mu}$, $V(U_n) = \boldsymbol{\lambda}'\boldsymbol{\Sigma}\boldsymbol{\lambda}$ とする．中心極限定理（リンドバーグ＝レヴィ）を U_n に応用すると $\sqrt{n}(\bar{U}_n - \boldsymbol{\lambda}'\boldsymbol{\mu})$ の漸近分布は $N_1(0, \boldsymbol{\lambda}'\boldsymbol{\Sigma}\boldsymbol{\lambda})$ である．この結果は $\bar{U}_n = \boldsymbol{\lambda}'\bar{\mathbf{U}}_n$ ということから求められる．

他の極限定理についてはこの章の終りの問題 **4. 1** から **4. 17** を見よ．

2c. 6　独立確率変数の和

$X_1, X_2, \cdots$ を独立な確率変数の列とし，部分和 $X_1 + \cdots + X_n$ を S_n であらわす．大数の法則（**2c. 3** 参照）と中心極限定理（**2c. 5** 参照）は，(a_n, b_n) を適当に定義された一対の数値の系列とすると，$(S_n - a_n)/b_n$ の行動に関係している．これについての研究をさらに進め，統計的応用に有用な二，三の一般的な定理を引用しておこう．証明は省略するが興味をもつ読者は Loève (1955) を調べればよい．

（i）**コルモゴロフの定理**　$E(X_n) = 0$, $E(X_n^2) = \sigma_n^2 < \infty$, $b_n \uparrow \infty$ とする．もし

$$\sum_1^\infty (\sigma_n^2/b_n^2) < \infty$$

ならば

$$\frac{S_n}{b_n} \xrightarrow{\text{a.s.}} 0 \qquad (2c.\ 6.\ 1)$$

である．

たとえば，もし $X_1, X_2, \cdots$ が $E(X_1) = 0$, $V(X_1) = \sigma^2 < \infty$ であるような独立で同一の分布に従うとすると，$S_n/\sqrt{n}\log n \xrightarrow{\text{a.s.}} 0$ である．

（ii）**重複対数の法則**　$X_1, X_2, \cdots$ が $E(X_1) = 0$, $V(X_1) = \sigma^2 < \infty$ であるような確率変数で独立で同一の分布に従うとする．さらに $h(n) = \sigma(2n \log\log n)^{1/2}$ とする．このとき

$$\limsup[S_n/h(n)] = 1 \quad \text{ほとんど確実に} \qquad (2c.\ 6.\ 2)$$

$$\liminf[S_n/h(n)] = -1 \quad \text{ほとんど確実に} \qquad (2c.\ 6.\ 3)$$

である．

(2c. 6. 2) と (2c. 6. 3) とはすべての $\varepsilon > 0$ にたいしてそれぞれ次のことをあらわしている：

$$P(\text{無限回も } |S_n| \geqq (1-\varepsilon)h(n) \text{ となること}) = 1, \qquad (2c.\ 6.\ 4)$$

$$P(\text{無限回も } |S_n| \geqq (1+\varepsilon)h(n) \text{ となること}) = 0. \qquad (2c.\ 6.\ 5)$$

重複対数の法則から任意の大きな k にたいして

$$P(\text{無限回も } S_n \geqq k\sqrt{n} \text{ となること}) = 1 \qquad (2c.\ 6.\ 6)$$

であることが直ちに示される．これは十分大きい n にたいして $(1-\varepsilon)h(n) > k\sqrt{n}$ だからである．

（iii）$X_1, X_2, \cdots$ を $E(X_i) = 0$, $V(X_i) = \sigma_i^2 < \infty$ であるような独立な確率変数とする．さらに $s_n^2 = \sigma_1^2 + \cdots + \sigma_n^2$ とする．もし $s_n^2 \to \infty$, $|X_n|/s_n = o(s_n^2/\sqrt{\log\log s_n})$, $t_n = (2\log\log s_n^2)^{1/2}$ ならば

$$\limsup \frac{S_n}{t_n s_n} = 1 \quad \text{ほとんど確実に} \qquad (2c.\ 6.\ 7)$$

$$\liminf \frac{S_n}{t_n s_n} = -1 \quad \text{ほとんど確実に} \qquad (2c.\ 6.\ 8)$$

である．

証明については Loève (1955) を参照されたい．

2d　確率測度の族と統計学の問題

2d. 1　確率変数の族

実際の場合には実験の結果に適用できる正確な確率測度 P は知られていない．しかしその確率測度はある測度族の要素として，集合 Θ に属するパラメータ θ $(\theta \in \Theta)$ によって指定されるような P_θ として考えることはできよう．統計学の問題は，観察事象（結果を示す標本空間の1点に対応する）に基づいて未知のパラメータ θ について有用な判定を行うことを取り扱っている．

確率測度の族をもつ標本空間は $(\Omega, \mathscr{B}, P_\theta, \theta \in \Theta)$ で示される．

2d. 2　十分統計量の概念

統計量 T（確率変数）が次に示す同値な条件のどれか1つをみたすとき，またそのときにかぎって，T は測度の族 P_θ にたいして十分であるという．

（a）$P(A|t)$ はすべての $A \in \mathscr{B}$ にたいして θ に無関係である．

（b）$E(Y|t)$ は $E(Y)$ が存在するようなすべての確率変数 Y にたいして θ に無関係である．

（c）すべての確率変数 Y の $T = t$ のときの条件つき分布は，いつも存在し，θ と無

関係である．

(b) が (a) を含むことはAの特性関数としてYを選べば明らかである．(a) が (c) を含むことは明らかで，また (c) が (b) を含むことも明らかである．したがってすべての3つの条件が同値である．

$P(A|t)$ と $E(Y|t)$ の定義については **2b. 3** を参照し，条件つき分布関数については **2a. 8** を参照されたい．

2d. 3　十分性の特徴づけ

十分統計量に関する **2d. 2** で与えた基準は個々の場合には適用するのが難かしい．確率密度関数がもし存在するならば，その形によってその統計量が十分であるかどうかを調べることができる．

（i）標本空間を離散型とし，$P_\theta(\omega)$ を点集合 $\{\omega\}$ の確率とする．θ にたいしてTが十分統計量である必要十分条件は，次のような分解

$$P_\theta(\omega) = g_\theta[T(\omega)]h(\omega) \tag{2d. 3. 1}$$

ができることである．ここで第1因子は θ に依存し，ω への依存は $T(\omega)$ を通してのみであるが，第2の因子はθに無関係である．

(2d. 3. 1) が成り立つと仮定する．このとき

$$P_\theta(T=t) = \sum_{T(\omega)=t} P_\theta(\omega) = g_\theta(t) \sum_{T(\omega)=t} h(\omega)$$
$$= g_\theta(t)H(t).$$

$T=t$ のときの ω の条件つき確率は

$$\frac{P_\theta(\omega)}{P_\theta(t)} = \frac{g_\theta(t)h(\omega)}{g_\theta(t)H(t)} = \frac{h(\omega)}{H(t)}$$

で，これはθに無関係である．このことからTは十分統計量である．

必要性は，もし

$$\frac{P_\theta(\omega)}{P_\theta(t)} = k(\omega, t)$$

がθに無関係であるとすると

$$P_\theta(\omega) = P_\theta(t)k(\omega, t)$$

であることから明らかである．

（ii）一般分解定理　P_θ を確率測度の族とする．また，σ 有限測度 μ に関して P_θ から確率密度 $p_\theta = dP_\theta/d\mu$ が導出されるとする．このとき非負可測関数 $g_\theta[T(\omega)]$ と $h(\omega)$ とが存在し

$$P_\theta(\omega) = g_\theta[T(\omega)]h(\omega) \tag{2d. 3. 2}$$

であるとき，またそのときにかぎってTは P_θ にたいして（または単に θ にたいして）十分である．

(2d. 3. 2) の形の分解をもっとも一般的に証明するには測度理論のいくつかのより高度の定理がいる (Halmos and Savage (1949), Bahadur (1954) を参照)．標本空間が R^n で，Tが連続な1階の偏導関数をもつときには，(2b. 3. 12) で展開したようにTが定数となる表面の上での条件つき密度の概念を用いれば証明は容易である．

十分性とフズルバザールの予想　$X_1, \cdots, X_n$ を分布関数 F_θ, $\theta \in \Theta$ をもつ独立な同一分布に従う確率変数とする．定義に従って，もし，Tが与えられたときの任意の他の統計量Yの条件つき分布が θ に無関係のとき可測関数 $T = f(X_1, \cdots, X_n)$ は十分統計量である．フズルバザールはTの十分性を確かめるには単に周辺の十分性を調べれば十分であること，すなわち，Tが与えられたときの X_i の条件つき分布が任意の i にたいして θ に無関係であることを調べればよいと予想した．この定理の部分的な証明は J. K. ゴーシュによって与えられた．また，より一般的な証明は V. N. Sudakov (1971) によって与えられた．

付録 2A　スティルチェス積分，ルベーグ積分

$F(x)$ は有界変動で分布関数のように左連続な関数とする．有限な区間 (a, b) と関数 $f(x)$ が与えられたとき，次のような和

$$J_n = \sum_{i=1}^{n} f(x_i')[F(x_i) - F(x_{i-1})] \tag{2A. 1}$$

をつくることができる．ここで，点 x_i は (a, b) を適当に $a < x_1 < \cdots < x_n < b$ のように分割するものとし，任意の $x_i' \in (x_{i-1}, x_i)$ とする．ここでリーマン積分では $F(x_i) - F(x_{i-1})$ のかわりに区間の長さ $(x_i - x_{i-1})$ を用いて同じような和が考えられていることを注意しよう．もし，(2A. 1) が各区間の長さを0にするように $n \to \infty$ とするとき

$$J = \lim_{n\to\infty} J_n$$

であるならば，Jを $f(x)$ の $F(x)$ に関するスティルチェス積分といい

$$J = \int_a^b f(x)dF(x)$$

とあらわす．広義の積分（変格積分）を次式

$$\lim_{\substack{a\to-\infty \\ b\to\infty}} \int_a^b f(x)dF(x) = f(x)dF(x)$$

で定義する．

リーマン積分と異なる1つの点は，端点が積分に含まれているかいないかを指定する必要があるということである．定義から次のことは容易に示される．

$$\int_a^b f(x)dF(x) = \int_{a+0}^b f(x)dF(x) + f(a)[F(a+0) - F(a)]$$

ここで $a+0$ は端点 a が積分に含まれないことを示す．もし $F(x)$ が点 a で飛躍をもつならば

$$\int_a^b f(x)dF(x) - \int_{a+0}^b f(x)dF(x) = f(a)[F(a+0) - F(a-0)]$$

となるから，長さがゼロになるような区間上での積分が必ずしもゼロとはならない．以下区間の下端は必ず含まれるが，上端は必ずしも含まれないと約束する．この約束により

$$\int_a^b dF(x) = F(b) - F(a)$$

である．

いま，

$$F(x) = \int_{-\infty}^x p(x)dx$$

となるような関数 $p(x)$ があるとすると，スティルチェス積分はリーマン積分に帰着し，

$$\int f(x)dF(x) = \int f(x)p(x)dx$$

である．

次に有限の測度空間におけるさらに一般的な積分を定義する．

$(\Omega, \mathscr{B}, \mu)$ を測度空間とし，s を単純可測関数とする．すなわち，$s = \sum_1^n \alpha_i \chi_{A_i}$ で，ここに $\alpha_1, \cdots, \alpha_n$ は s の相異なる値，χ_{A_i} は集合 $A_i \in \mathscr{B}$ の定義関数とする．このとき，Ω の上での s の積分を

$$\int_\Omega s d\mu = \sum_1^n \alpha_i \mu(A_i) = \int s d\mu$$

と定義する．ここで積分領域が了解されているときはそれを省略することもある．また，$\alpha_i = 0$，$\mu(A_i) = \infty$ のときには $0 \cdot \infty = 0$ と約束する．いま f は非負の可測関数とする．このとき，f のルベーグ積分を

$$f d\mu = \sup \int s d\mu$$

と定義する．ここで上限は $0 \leqslant s \leqslant f$ であるようなすべての単純可測関数 s についてとるものとする．

もし一般の関数 f が

$$\int |f| d\mu < \infty$$

であるならば，f はルベーグ積分可能 $[L_1(\mu)]$ という．この場合 f のルベーグ積分は

$$\int f^+ d\mu - \int f^- d\mu$$

であり，ここに f^+ と f^- はそれぞれ f の正および負の部分である．

上式でいずれか1つの積分が有限であるなら，可測関数 f の積分の値域を区間 $[-\infty, \infty]$ として，

$$f d\mu = \int f^+ d\mu - \int f^- d\mu$$

により定義したいことがある．このとき積分の値は，$[-\infty, \infty]$ の数値となる．

特に，μ が分布関数 F によって生成される場合には，ルベーグ積分はスティルチェス・ルベーグ積分と呼ばれる．

付録 2B　測度論と積分における重要な定理

ここでは測度論，積分論における多くの結果をまとめておく．証明については測度論，積分論の教科書を参照されたい．これらの結果を知っていることは，高度の理論統計学の学習および研究において有用である．

さて以下で可測空間 $(\Omega, \mathscr{B}, \mu)$ の上で定義された可測関数 g, g_n，$n = 1, 2, \cdots$ を考え，収束の型としては次のものを考える．

(a)　$g_n \to g$，a. e. μ（測度 μ に関してほとんどいたるところで収束）：零集合以外の各点で $g_n \to g$ のとき

(b)　$g_n \xrightarrow{w} g$（弱収束）：すべての可積分関数 f について $\int g_n f d\mu \to \int g f d\mu$ のとき

(c)　$g_n \xrightarrow{L_p} g$（p 次平均収束）：$\int |g_n - g|^p d\mu \to 0$ のとき，$p = 1$ のときは平均収束という．

(d)　$g_n \xrightarrow{\mu} g$（測度収束）：すべての $\varepsilon > 0$ について $\lim_{n\to\infty} \mu\{|g_n - g| \geqslant \varepsilon\} = 0$ のとき

(d) の定義はほとんどいたるところ有界な関数 g_n に限られる．

以下の補助定理は定義より容易に証明できる．

補助定理　2B. 1　(i)　$0 < r \leqslant p$ について $\mu(\Omega) < \infty$ のとき $g_n \xrightarrow{L_p} g \Rightarrow g_n \xrightarrow{L_r} g$

(ii)　$\int |g_n - g| d\mu \to 0 \Leftrightarrow n \to \infty$ のときすべての $A \in \mathscr{B}$ について一様に

$$\int_A g_n d\mu \to \int_A g d\mu,$$

ここで $\{g_n, g\} \subset L_1$．さらに $g_n \xrightarrow{L_1} g \Rightarrow g_n \xrightarrow{\mu} g$.

補助定理　2B. 2　一様有界な関数列 g_n（あるいは可積分関数のみからなる列）は次の場合にかぎり有界関数（あるいは可積分関数）g に弱い意味で収束する：

$$\int_A g_n d\mu \to \int_A g d\mu \tag{2B. 1}$$

ここに A は $\mu(A) < \infty$ であるようなすべての A.

補助定理　2B. 3　g_n を可測関数列とし，μ についてほとんどいたるところで(a.e. μ) $g_n \to g$ であり，ほとんどいたるところで $g_n \geqslant f$（あるいは $g_n \leqslant h$）とする．ここに f（あるいは h）は可積分とする．g が可積分ならば

$$\int g_n d\mu \to \int g d\mu \Rightarrow \int |g_n - g| d\mu \to 0. \tag{2B. 2}$$

特に p, p_n，$n = 1, 2, \cdots$ が μ に関する確率密度関数で $p_n \to p$ (a. e. μ) ならば，$\int |p_n - p| d\mu \to 0$.

ルベーグの単調収束定理　g_n を非負で，非減少の可測関数列とし，μ についてほとんどいたるところ $g_n \to g$ とする．このとき g は可測で

$$\int g_n d\mu \to \int g d\mu. \tag{2B. 3}$$

この定理の結果として，次の系をうる．

系 1.　積分は非負関数族の上で σ 加法的である．いいかえればもし g_n が非負の関数ならば

$$(\sum g_n)d\mu = \sum \int g_n d\mu \tag{2B. 4}$$

である．

系 2.（ファトゥーの補助定理）　$g_n \geqslant f$（可積分）を可積分関数列とする．このとき $\liminf g_n$ は可積分で

$$\int \liminf_{n\to\infty} g_n d\mu \leqslant \liminf_{n\to\infty} \int g_n d\mu \tag{2B. 5}$$

である．

系 3.　g を非負とする．このとき

（a）　$\mu(E) = 0$ ならば

$$\int_E g d\mu = 0. \tag{2B. 6}$$

（b）

$$\lambda(E) = \int_E g d\mu, \qquad E \in \mathscr{B} \tag{2B. 7}$$

は $(\Omega, \mathscr{B})$ 上の測度である．

（c）　$\int g d\mu = 0$ ならば，かつそのときにかぎりほとんどいたるところで $g = 0$.　(2B. 8)

系 4.　$g_n \leqslant h$ を可積分関数列とするとき

$$\limsup_{n\to\infty} \int g_n d\mu \leqslant \int \limsup_{n\to\infty} g_n d\mu. \tag{2B. 9}$$

系 5.　f と F は可積分とし $f \leqslant g_n \leqslant F$ で，$g_n \to g$ とする．このとき

$$\int g_n d\mu \to \int g d\mu. \tag{2B. 10}$$

ルベーグの上限収束定理　g_n をすべての n, ω について $|g_n(\omega)| \leqslant G(\omega)$ であるような可測関数列とし，G は測度 μ について有限または無限の集合 S の上で可積分とする．いま，$g_n \to g$ が S 上で a. e. μ であるか，または単に測度収束するならば，

$$\int_S g_n d\mu \to \int_S g d\mu. \tag{2B. 11}$$

である．実は，このとき

$$\int_S |g_n - g| d\mu \to 0, \qquad \text{すなわち，} A \text{について一様に} \int_A g_n d\mu \to \int_A g d\mu \tag{2B. 12}$$

である．（これは $\{g_n\} \subset L_1$ が一様に可積分であり，$g_n \to g$ が測度収束でであるというのと同等である．）

上の定理の結果として次の系を得る．

系 1.　各 $\theta \in R$ について $g(\omega, \theta)$ を $\omega \in \Omega$ の可測関数とし，$\theta \to \theta_0$ のとき $g(\omega, \theta) \to g(\omega, \theta_0)$ とする．さらに θ_0 を含む開区間上のすべての ω と θ について $|g(\omega, \theta)| \leqslant G(\omega)$ とし，G は S の上で可積分とする．このとき $\theta \to \theta_0$ で

$$\int_S g(\omega, \theta) d\mu \to \int_S g(\omega, \theta_0) d\mu \tag{2B. 13}$$

である．

系 2.　$dg(\omega, \theta)/d\theta$ が θ の区間 (a, b) で存在し，$|dg(\omega, \theta)/d\theta| < G(\omega)$ なる $G(\omega)$ が可積分とするなら，(a, b) で

$$\frac{d}{d\theta} \int g(\omega, \theta) d\mu = \int \frac{d}{d\theta} g(\omega, \theta) d\mu \tag{2B. 14}$$

である．

拡張したルベーグの上限収束定理（Pratt (1960)）　いま

（a）　$f_n \leqslant g_n \leqslant F_n$

（b）　μ についてほとんどいたるところで次の条件が成り立つ：

$$f_n \to f, \quad g_n \to g, \quad F_n \to F$$

（c）　f_n, F_n, f, F は可積分で

$$\int f_n d\mu \to \int f d\mu, \qquad \int F_n d\mu \to \int F d\mu$$

とする．このとき $\int g d\mu$ は有限で $\int g_n d\mu \to \int g d\mu$ である．さらに

$$\int |f_n - f| d\mu \to 0 \text{ ならば } \int |g_n - g| d\mu \to 0 \tag{2B. 15}$$

である．特に μ についてほとんどいたるところで $f_n \geqslant h$ あるいは $f_n \leqslant k$ ならば $\int |g_n - g| d\mu \to 0$ である．

この定理より，次の系をうる．

系 1.　θ を実補助変数とし，D を θ に関する微分，$\int$ を μ に関する積分とする．また，θ_0 を区間 I の内点とし，

（i）　すべての $\theta \in I$ について $Dg(\omega, \theta)$，$Df(\omega, \theta)$，$DF(\omega, \theta)$ が存在する．

（ii）　すべての $\theta \in I$ について $Df(\omega, \theta) \leqslant Dg(\omega, \theta) \leqslant DF(\omega, \theta)$

（iii）　$h = f$，F でその値はすべて有限とし，$D\int h(\omega, \theta_0) d\mu = \int Dh(\omega, \theta_0) d\mu$ ならば，

$$D\int g(\omega, \theta_0) d\mu = \int Dg(\omega, \theta_0) d\mu \tag{2B. 16}$$

である．

フビニの定理　$(\Omega_i, \mathscr{B}_i, \mu_i)$ $i = 1, 2$ を2つの有限な測度空間とする．$\Omega_1 \times \Omega_2$ の上で定義された関数 $g(\omega_1, \omega_2)$ が $\mathscr{B}_1 \times \mathscr{B}_2$ について可測であり，かつ非負であるか，または測度 $\mu = \mu_1 \times \mu_2$ に関して $\Omega = \Omega_1 \times \Omega_2$ 上で可積分であるかのいずれかであるならば

$$\int_\Omega g d\mu = \int_{\Omega_1} d\mu_1 \int_{\Omega_2} g d\mu_2 = \int_{\Omega_2} d\mu_2 \int_{\Omega_1} g d\mu_1 \tag{2B. 17}$$

である．

ラドン・ニコディムの定理　μ と ν を $(\Omega, \mathscr{B})$ の上の σ 有限な測度とする．このとき可測関数 g が存在し

$$\nu(A) = \int_A g d\mu \qquad \text{すべての } A \in \mathscr{B} \text{ について} \tag{2B. 18}$$

となるのは，ν が μ について絶対連続であるとき（すなわち，$\mu(A) = 0$ なるすべての A

$\in \mathscr{B}$ について $\nu(A)=0$ であるときであり，この関係を $\nu << \mu$ と書く）であり，かつ，そのときにかぎる．

測度は $\mu(\Omega)<\infty$ ならば有限とよび，各 i について $\bigcup A_i=\Omega$ で $\mu(A_i)<\infty$ であるような互いに素な集合 $A_1, A_2, \cdots$ が存在するならば σ 有限であるということを注意しておく．

これらの定理の証明については Halmos (1950) および Munroe (1953) を参照されたい．測度論および関連理論の入門が Cramer (1946) に含まれているが，このなかではとくに確率との関係にふれている．

ボレル＝カンテリの補助定理　$\{A_n\}$，$n=1,2,\cdots$ を集合の系列とし，Aをこれらの集合の無限個に共通に含まれる要素の集合，Pを確率集合関数とする．このとき

（a）　$\sum P(A_n)<\infty \Rightarrow P(A)=0$

（b）　$\sum P(A_n)=\infty$ で，事象 A_n が独立 $\Rightarrow P(A)=1$.

である．

集合Aを

$$A=\bigcap_{r=1}^{\infty}\bigcup_{n=r}^{\infty}A_n$$

と指定できることは容易にわかる．このことから，すべてのrについて

$$A\subset\bigcup_{n=r}^{\infty}A_n$$

であり，$\sum P(A_n)$ は収束するから，十分大きなrについて

$$P(A)\leqslant P(\bigcup_{r}^{\infty}A_n)\leqslant\sum_{r}^{\infty}P(A_n)<\varepsilon$$

である．すなわち $P(A)=0$ であって (a) が証明された．

A^c をAの補集合，すなわち $\{A_n\}$ の有限個の集合にのみ共通な要素の集合とする．A^c は有限個の $\{A_n^c\}$ の集合に含まれている点の集合である．したがって

$$A^c=\bigcup_{r=1}^{\infty}\bigcap_{n=r}^{\infty}A_n^c$$

$$1-P(A)=P(A^c)=P(\bigcup_{r=1}^{\infty}\bigcap_{n=r}^{\infty}A_n^c)\leqslant\sum_{r=1}^{\infty}P(\bigcap_{n=r}^{\infty}A_n^c)$$

事象の独立性を使うと

$$=\sum_{r=1}^{\infty}\prod_{n=r}^{\infty}[1-P(A_n)].$$

$\sum P(A_n)=\infty$ であるから，各rについて無限積はゼロに発散する．したがって，$P(A)=1$ となり，(b) は証明された．

付録 2C　不変性

確率空間族の不変性　2d の場合と同じように次のような確率空間の族

$$(\Omega, \mathscr{B}, P_\theta, \theta\in\Theta) \tag{2C.1}$$

を考える．gをΩからΩの上への1対1変換としよう．また，シグマ集合体$\{gB : B\in\mathscr{B}\}$

を $g\mathscr{B}$ であらわす．与えられた測度 P_θ に対応してすべての $B\in\mathscr{B}$ にたいして $gP_\theta(gB)=P_\theta(B)$ という関係をみたすような gP_θ を $g\mathscr{B}$ の上に導かれた測度とする．Ω の上の任意の関数 ϕ は $g\phi(g\omega)=\phi(\omega)$，$\omega\in\Omega$ であるような新しい関数 $g\phi$ を生成する．そのときgは $(\Omega, \mathscr{B}, P_\theta)$ と $(\Omega, g\mathscr{B}, gP_\theta)$ との間の同型写像を生みだす．

Gを変換群とする．このとき $(\Omega, \mathscr{B}, P_\theta, \theta\in\Theta)$ は，もしいかなる $g\in G$ についても

$$g\Omega=\Omega,\quad g\mathscr{B}=\mathscr{B},\quad gP_\theta=P_{\theta'},\quad \theta,\theta'\in\Theta \tag{2C.2}$$

であるならばGのもとで不変であるという．この (2C.2) の最後の条件から g によって定まる θ' を $\bar{g}\theta$ であらわす．

このとき，$\bar{G}=\{\bar{g} : g\in G\}$ は Θ からΘの上への1対1の写像群であること，$\bar{G}$はGの準同型写像の像であること（$\bar{G}$ とGは必ずしも同型とは限らない）が示される (Ferguson (1967))．

統計的手続きの不変性　実際の場では，われわれは事象を観測するが（すなわちΩのなかの点をきめうるが），θ の真の値や P_θ，$\theta\in\Theta$ の族のなかのどの確率測度に属するかなどはわからない．統計的問題とは観測事象のもとで未知のθについて1つの陳述を行うことである．いま可能なすべての陳述（あるいは決定）の族をDとする（これを決定空間という）．このとき，ΩからDのなかへの写像である関数 δ（決定関数という）が必要となる．θ が真の値であったときに決定 $d\in D$ を行うときにひき起こされる損失を $L(\theta, d)$ であらわそう．最適な統計的手続きとは，ある意味において損失を最小にするような決定関数 δ を用いることである．

Gを変換群とし，この変換にたいして $(\Omega, B, P_\theta, \theta\in\Theta)$ は不変であるとする．各 $g\in G$ にたいして，決定空間Dの自分の上への次のような変換 g^* が存在するとしよう．すなわち，g^* は準同型，つまり

$$(g_1g_2)^*=g_1^*g_2^*$$

であり，また，変換を行っても損失関数が不変である，つまり

$$L(\bar{g}\theta, g^*d)=L(\theta, d),\quad \theta\in\Theta,\quad d\in D \tag{2C.3}$$

である．このとき統計的手続きあるいは決定関数 δ は，もし

$$\text{すべての } \omega\in\Omega,\quad g\in G \text{ について，} \delta(g\omega)=g^*\delta(\omega) \tag{2C.4}$$

ならば，Gのもとで不変という．

付録 2D　統計量，部分空間，十分性

$(\Omega, \mathscr{B})$ を任意の可測空間，R^k をk（有限）次元のユークリッド空間，$\mathscr{X}$ を R^k 上のボレル集合体とする．

定義 1.　2d でのように，統計量TはΩから R^k への関数で，すべての $X\in\mathscr{X}$ にたいして，$T^{-1}(X)\in\mathscr{B}$ であるものと定義される．

Ω上で定義された R^k への任意の関数Tにたいして，集合 $\{T^{-1}(X) : X\in\mathscr{X}\}$ の全体を$\mathscr{B}_T$とあらわす．このときTが統計量(1)，つまり定義1による統計量であるのは，$\mathscr{B}_T\subset\mathscr{B}$ であるとき，そのときにかぎる．

部分σ集合体$\mathscr{C}\subset\mathscr{B}$は，$\mathscr{C}=\mathscr{B}_T$ ならば，統計量Tによって導かれたという．次の命題 (i) の示すように，部分σ集合体のすべてが統計量 (1) によって導かれるわけではない．

（i）部分 σ 集合体 $\mathscr{C} \subset \mathscr{B}$ は，$\mathscr{C}$ が可算的に生成されるとき，そのときにかぎって統計量 (1) によって導かれる．

そのときにかぎりというのは，R^k の $\mathscr{X}$ が可分であり，統計量 (1) である T から導かれる部分集合体 $\mathscr{B}_T$ が可分であることによる．

逆を証明するには，$\mathscr{C} \subset \mathscr{B}$ を可算的に生成されるものとし，Ω から実直線 R のなかへの関数 T を，

$$T(\omega) = \sum_{n=1}^{\infty} \frac{2}{3^n} I_{C_n}(\omega)$$

とする．ここで I は定義関数とし，$\{C_n,\ n \geqq 1\}$ は $\mathscr{C}$ を生成する可算な族とする．T が統計量 (1) であり $\mathscr{B}_T \subseteq \mathscr{C}$ であることは容易にわかる．逆の包含関係を示すために，$n \geqq 1$ とし，B_n を3進数であらわしたとき2が n 桁めにあるような $[0, 1]$ のなかの実数の集合とする．このとき，(a) B_n はボレル集合であり，(b) $T^{-1}(B_n) = C_n$ であることは容易にわかる．したがって，$\{C_n, n \geqq 1\} \subseteq \mathscr{B}_T$ で $\mathscr{C} = \mathscr{B}_T$ となり，命題が証明された．

定義 2.　Bahadur (1954, 1955) は，統計量 T を Ω 上で定義された任意の関数で，その値域空間は任意であり（ユークリッド空間とはかぎらない），可測性の条件もつけないものとして呼ぶ．バハデュールの定義に従う統計量を統計量 (2) と定義した．

統計量 (2) の T によって導かれる部分 σ 集合体の概念は次のように定義される．$\mathscr{Y}_T$ を集合 $\{A \subseteq Y : T^{-1}(A) \in \mathscr{B}\}$ とする．ここで Y は T の正確な値域，すなわち $Y = T(\Omega)$ とする．次のことは明らかである．(a) $\mathscr{Y}_T$ は Y の上の σ 集合体である．(b) T は $(\Omega, \mathscr{B})$ から $(Y, \mathscr{Y}_T)$ への可測変換である．次に $\mathscr{B}_T$ を集合 $\{T^{-1}(B) : B \in \mathscr{Y}_T\}$ とする．このとき $\mathscr{B}_T$ は $\mathscr{B}$ の部分 σ 集合体である．

部分 σ 集合体 $\mathscr{C} \subset \mathscr{B}$ は，もし関数 T があり $\mathscr{C} = \mathscr{B}_T$ ならば，統計量 (2) から導かれるという．すべての $\mathscr{B}$ の部分 σ 集合体は統計量 (2) から導かれるのか，あるいはさらにすべての可算的に生成される部分 σ 集合体は統計量 (2) から導かれるのかという問題が生ずる．この答は，以下にブラックウェルの結果の示すように否定的である．

（ii）$\mathscr{C}$ を $\mathscr{B}$ の真の部分 σ 集合体とし，すべての $\omega \in \Omega$ にたいして $\{\omega\} \in \mathscr{C}$ とする．このとき $\mathscr{C}$ は統計量 (2) からは導かれない．

いま $\mathscr{C}$ は統計量 T から導かれ $\mathscr{B}_T = \mathscr{C}$ とする．$\{\omega\} \in \mathscr{C}$ であるから，$\{\omega\} \in \mathscr{B}_T$ であり，したがって $\{\omega\} = T^{-1}(B)$ であるような $B \in \mathscr{Y}_T$ が存在する．B が T の正確な定義域に含まれるので，容易に $B = \{T(\omega)\}$ が導かれる．したがって，すべての $\omega \in \Omega$ について，$\{\omega\} = T^{-1}\{T(\omega)\}$ である．いいかえると，T は Ω 上で定義された Y の上への1対1関数で，任意の $A \subset \Omega$ について $A = T^{-1}(T(A))$ である．いま $A \in \mathscr{B}$ とする．上の論議より，$A \subset \mathscr{B}_T$ すなわち $A \in \mathscr{C}$ が導かれる．それゆえ $\mathscr{B} = \mathscr{C}$ となり，これは矛盾である．

Bahadur and Lehmann (1955) は統計量 (2) から導かれる $\mathscr{B}$ の部分 σ 集合体の特徴づけを行った．

われわれは統計量 (1) と (2) とから導かれる $\mathscr{B}$ の部分 σ 集合体の概念をもつに到った．次にこのいずれの一方の概念にも他の概念が含まれないことを示そう．

Ω を実数軸とし，$\mathscr{B}$ を実数軸の可算な部分集合およびその補集合の σ 集合体とし，$\mathscr{C} = \mathscr{B}$ とする．T を Ω の上の恒等写像とすると，$\mathscr{C}$ が統計量 (2) から導かれることがわかる．しかし $\mathscr{C}$ は可算的には生成されず，統計量 (1) からは導かれない．

いま Ω を実数軸とし，M を実数軸上の一定の，ただしボレル集合ではない任意の部分集合とし，$\mathscr{B}$ を M と実軸上のすべてのボレル集合とを含む Ω の上の最小の部分 σ 集合体とする．$\mathscr{C}$ を実軸のボレル σ 集合体とする．このとき $\mathscr{C}$ は可算的に生成されるので統計量 (1) から導かれる．しかし $\mathscr{C}$ はすべての孤立点を含みかつ $\mathscr{B}$ の完全な部分 σ 集合体であるので，ブラックウェルの命題から $\mathscr{C}$ は統計量 (2) からは導かれない．

部分集合体に関する条件つき期待値　任意の部分 σ 集合体 $\mathscr{B}_0 \subset \mathscr{B}$ を考える．ϕ が Ω 上で定義された任意の $(\mathscr{B}, P)$ 積分可能な関数であるとすると，$\mathscr{B}_0$ を与えたときの ϕ の条件つき期待値は，次の条件をみたす任意の $\mathscr{B}_0$ 可測関数 $E_P(\phi|\mathscr{B}_0)$ によって定義される．すなわち，すべての $B_0 \in \mathscr{B}_0$ について

$$\int_{B_0} E_P(\phi|\mathscr{B}_0)dP = \int_{B_0} \phi dP. \tag{2D. 1}$$

ラドン＝ニコディムの定理によって $E_P(\phi|\mathscr{B}_0)$ の存在と，$\mathscr{B}_0$ 可測な関数の同等性からその一意性，すなわち (2D. 1) をみたす他の関数があっても，その違いは P 零集合の上においてのみであるということが確かめられる．

もし Ω の上のすべての $(\mathscr{B}, P_\theta)$ 可積分関数 ϕ にたいして，Ω の上の $\mathscr{B}_0$ 可測な関数 ϕ_0 が存在し，すべての $\theta \in \Theta$ について

$$E_\theta(\phi|\mathscr{B}_0)$$

と同等であるならば，部分集合体 $\mathscr{B}_0$ は，族 $(\Omega, \mathscr{B}, P_\theta, \theta \in \Theta)$ にたいして十分であるという．

さらに部分集合体，十分部分集合体について研究したい場合には，読者は Bahadur (1954, 1955) および Bahadur and Lehmann (1955) を参照されたい．

付録 2E　特性関数の非負定符号性

ボホナーの定理　R 上で定義された関数 $f(\cdot)$ が非負定符号で連続であり $f(0)=1$ であるのは，$f(\cdot)$ が特性関数であるときであり，そのときにかぎる．

［複素数値関数 f が非負定符号であるのは，任意の整数 n，実数 $t_1, \cdots, t_n$，および R 上の関数 $h(t)$ について

$$\sum_{i=1}^{n} \sum_{j=1}^{n} f(t_i - t_j) h(t_i) \overline{h(t_j)} \geqq 0 \tag{2E. 1}$$

が成り立つときである．］

十分性は証明容易である．必要性を証明するためには，(2E. 1) より，各 x および $\sigma^2 > 0$ にたいして

$$\iint f(u-v) \exp\{-i(u-v)x - (u^2+v^2)\sigma^2\} du dv \geqq 0 \tag{2E. 2}$$

であることに注目する．なぜなら (2E. 2) の積分は (2E. 1) の式の和によって近似できるからである．いま $t = u - v$，$s = u + v$ とおくと (2E. 2) で $u^2 + v^2 = (t^2 + s^2)/2$ であるから，これを s について積分して

$$g(x, \sigma) = \int f(t) \exp(-itx - t^2\sigma^2/2) dt/2\pi \geqq 0$$

が得られる．単調収束定理（付録 **2B** の (2B. 3)）より

$$\begin{aligned}\int g(x,\sigma)dx &= \lim_{\beta\to 0}\int g(x,\sigma)\exp(-x^2\beta^2/2)dx\\ &= \lim_{\beta\to 0}\iint f(t)\exp(-itx - t^2\sigma^2/2 - x^2\beta^2/2)dxdt/2\pi\\ &= \lim_{\beta\to 0}\int f(t)\exp(-t^2/2\beta^2 - t^2\sigma^2/2)dt/\beta\sqrt{2\pi}\\ &= \lim_{\beta\to 0}\int f(u\beta)\exp(-u^2/2 - u^2\beta^2\sigma^2/2)du/\sqrt{2\pi}\\ &= \int e^{-u^2/2}du/\sqrt{2\pi} = 1\end{aligned}$$

が得られる．すなわち $g(x,\sigma)$ は x の関数で各 σ についての密度関数である．同様な議論から，$f(t)$ の連続性を利用して

$$\begin{aligned}\int g(x,\sigma)\exp(itx)dx &= \lim_{\beta\to 0}\int g(x,\sigma)\exp(itx - x^2\beta^2/2)dx\\ &= \lim_{\beta\to 0}\int f(t+u\beta)\exp[-(t+u\beta^2/2) - u^2/2]du\\ &= f(t)\exp(-t^2\sigma^2/2)\end{aligned}$$

である．したがって各 σ について $f(t)\exp(-t^2\sigma^2/2)$ は特性関数である．そこで連続性の定理から

$$\lim_{\sigma\to 0} f(t)\exp(-t^2\sigma^2/2) = f(t)$$

は実際に特性関数である．この証明は Pathak (1966) による．

補足と問題

1　特性関数

1.1　分布関数が $F(x) = 1 - F(-x-0)$ ならば零のまわりで対称であるという．$F(x)$ が対称であるのは特性関数が実関数で偶関数のときであり，そのときにかぎることを示せ．

1.2　$f(t)$ が特性関数ならば $e^{f(t)-1}$ および $|f(t)|^2$ も特性関数である．

1.3　特性関数がそれぞれ

$$\cos t,\ (\cosh t)^{-1},\ \cos^2 t,\ (\cosh^2 t)^{-1}$$

であるような分布関数を求めよ．

1.4　すべての特性関数は

$$f(t) = c_1 f_d(t) + c_2 f_{ac}(t) + c_3 f_s(t)$$

の形に書け，$c_1 + c_2 + c_3 = 1$ で，$c_1 \geqslant 0$，$c_2 \geqslant 0$，$c_3 \geqslant 0$ であることを示せ．ただし，$f_d(t)$，$f_{ac}(t)$，$f_s(t)$ はそれぞれ（純）離散的，絶対連続および特異な分布である．

1.5　任意の特性関数 $f(t)$ について

$$1 - |f(2t)|^2 \leqslant 4[1 - |f(t)|^2]$$

であること，任意の実特性関数 $f(t)$ について

$$1 - f(2t) \leqslant 4[1 - f(t)],\qquad 1 + f(2t) \geqslant 2[f(t)]^2$$

であることを示せ．

1.6　X_1, X_2, X_3 が独立な確率変数で，$X_1 + X_2$ および $X_1 + X_3$ の分布が同一であるとする．このとき X_2 と X_3 は同じ分布であるか．

2　分布の積率　α_s および γ_s を確率変数 X の s 次の積率と s 次の絶対積率とする．

2.1　次の不等式を証明せよ：

(a)　$\gamma_{s-1}^2 \leqslant \gamma_{s-2}\gamma_s$

(b)　$\gamma_1 \leqslant \gamma_2^{1/2} \leqslant \cdots \leqslant \gamma_s^{1/s}$.

2.2　μ_s を s 次の中心積率とし $\beta_1 = \mu_3^2/\mu_2^3$，$\beta_2 = \mu_4/\mu_2^2$ と定義する．このとき $\beta_2 - \beta_1 - 1 \geqslant 0$ である．［ヒント：次の行列式の性質を利用せよ．

$$\begin{vmatrix} 1 & 0 & \mu_2\\ 0 & \mu_2 & \mu_3\\ \mu_2 & \mu_3 & \mu_4\end{vmatrix} \geqslant 0.\ ]$$

3　確率に関する不等式

3.1　X が確率変数で $a > 0$ にたいして $E(e^{aX})$ が存在するならば，$P(X \geqslant \varepsilon) \leqslant E(e^{aX})/e^{a\varepsilon}$ である．

3.2　コルモゴロフの不等式　$X_1, X_2, \cdots, X_n$ が独立な確率変数で $E(X_i) = 0$，$V(X_j) = \ = \sigma_i^2 < \infty$ とする．このとき

$$P\Big(\max_{1\leqslant k\leqslant n}\Big|\sum_1^k X_i\Big| > \varepsilon\Big) < \frac{1}{\varepsilon^2}\sum_1^n \sigma_i^2$$

を証明せよ．

3.3　ハイエク＝レニュイ(1955) の不等式　$X_1, X_2, \cdots$ を独立な確率変数とし $E(X_i) = 0$，$V(X_i) = \sigma_i^2 < \infty$ とする．いま $c_1, c_2, \cdots$ が非増加の正定数とするとき，任意の正の整数 m, n $(m < n)$ および $\varepsilon > 0$ について

$$P(\max_{m\leqslant k\leqslant n} c_k|X_1 + \cdots + X_k| \geqslant \varepsilon) \leqslant \frac{1}{\varepsilon^2}\Big(c_m^2\sum_1^m \sigma_i^2 + \sum_{m+1}^n c_i^2\sigma_i^2\Big)$$

である．この不等式を証明するには

$$Y = \sum_{k=m}^{n-1} S_k^2(c_k^2 - c_{k+1}^2) + c_n^2 S_n^2$$

を考えよ．ここに $S_k = X_1 + \cdots + X_k$ とする．このとき

$$E(Y) = c_m^2\sum_{k=1}^m \sigma_k^2 + \sum_{k=m+1}^n c_k^2\sigma_k^2$$

を示すことは容易である．いま，E_i $(i = m, m+1, \cdots, n)$ を $m \leqslant j < i$ について $c_j|S_j| < \varepsilon$ で，$c_i|S_i| \geqslant \varepsilon$ であるような事象とする．このとき事象 E_i は互いに排反で，

$$P\{\max_{m\leqslant k<n} c_k|S_k| \geqslant \varepsilon\} = \sum_{k=m}^n P(E_i)$$

である．

E_0 を $m \leqslant j \leqslant n$ について $c_j|S_j| < \varepsilon$ の事象とする．このとき条件つき期待値の定義により

$$E(Y) = \sum_{i=m}^n E(Y|E_i)P\{E_i\} + E[Y|E_0]P(E_0) \geqslant \sum_{i=m}^n E(Y|E_i)P\{E_i\}$$

である．また，$k \geqslant i$ について

$$\begin{aligned}E(S_k^2|E_i) &= E[\{S_i^2 + (X_{i+1} + \cdots + X_k)^2 + 2S_i(X_{i+1} + \cdots + X_k)\}|E_i]\\ &\geqslant E(S_i^2|E_i) + 2E[S_i(X_{i+1} + \cdots + X_k)|E_i]\end{aligned}$$

である．しかし，事象 E_i が起こる条件は変数 X_j の最初の i にのみ制限を課し，この条

件のもとで，これにつづく変数は互いに独立であり，S_i とも独立である．$j>i$ について $E(S_iX_j|E_i)=0$ であるから，不等式 $E(S_k^2|E_i)\geqslant E(S_i^2|E_i)$ が成り立つ．E_i が与えられるならば，$|S_i|\geqslant\varepsilon/c_i$ で $E(S_i^2|E_i)\geqslant\varepsilon^2/c_i^2$ となる．したがって $E(S_k^2|E_i)\geqslant\varepsilon^2/c_i^2$ が成り立つ．しかし

$$\begin{aligned}E(Y|E_i)&=\sum_{k=m}^{n-1}E(S_k^2|E_i)(c_k^2-c_{k+1}^2)+c_n^2E(S_n^2|E_i)\\&\geqslant\sum_{k=i}^{n-1}E(S_k^2|E_i)(c_k^2-c_{k+1}^2)+c_n^2E(S_n^2|E_i)\\&\geqslant\frac{\varepsilon^2}{c_i^2}\left[\sum_{k=i}^{n-1}(c_k^2-c_{k+1}^2)+c_n^2\right]=\varepsilon^2\end{aligned}$$

であり，したがって $E(Y)\geqslant\varepsilon^2\sum_{i=m}^{n}P(E_i)$ である．求める結果は $E(Y)$ として導いた式から出てくる．

問題3.2のコルモゴロフの不等式は，上で $m=1$，$c_1=c_2=\cdots=c_n=1$ とした特別の場合として得られる．

3.4　$X_1,\cdots,X_m$ を m 個の独立変数とし，$i=1,\cdots,m$ について $P(|X_i+\cdots+X_m|>\varepsilon)\leqslant\alpha$ とする．このとき

$$P\{\max_{1\leqslant i<m}|X_1+\cdots+X_i|>2\varepsilon\}\leqslant\frac{P(|X_1+\cdots+X_m|>\varepsilon)}{1-\alpha}$$

である．

3.5　チェビシェフ型の不等式　X を確率変数とし $E(X)=\mu$，$E(X-\mu)^2=\sigma^2$，$E(|X-\mu|^r)=\beta_r$，$E(|X-\mu|)=\nu$，$\delta=\nu/\sigma$ とする．このとき以下のことを示せ．

(a)　$P\left(\dfrac{|X-\mu|}{\beta_r^{1/r}}\geqslant\lambda\right)\leqslant\dfrac{1}{\lambda^r}$

(b)　$P\left(\dfrac{|X-\mu|}{\sigma}\geqslant\lambda\right)\leqslant\dfrac{\beta_r}{\sigma^r\lambda^r}$

(c)　$P(|X-\mu|\geqslant\lambda)\leqslant\dfrac{\beta_r}{\lambda^r}$

(d)　カンテリの不等式

$$P(X-\mu\leqslant\lambda)\leqslant\frac{\sigma^2}{\sigma^2+\lambda^2},\qquad\lambda<0$$

$$\geqslant1-\frac{\sigma^2}{\sigma^2+\lambda^2},\qquad\lambda\geqslant0.$$

(e)　ビークの不等式

$$P(|X-\mu|\geqslant\lambda\sigma)<\frac{1-\delta^2}{\lambda^2-2\lambda\delta+1},\qquad\lambda\geqslant\delta.$$

(f)　カンプ＝メイデルの不等式　X の分布は μ_0 のところで単峰性をもつとする．$\tau=\sigma^2+(\mu-\mu_0)^2$，$s=|\mu-\mu_0|/\sigma$ とするならば

$$P(|X-\mu_0|\geqslant\lambda\tau)\leqslant\begin{cases}1-\dfrac{\lambda}{\sqrt{3}},&\lambda\leqslant\dfrac{2}{\sqrt{3}}\\[2ex]\dfrac{4\lambda^2}{9},&\lambda\geqslant\dfrac{2}{\sqrt{3}}\end{cases}$$

$$P(|X-\mu|\geqslant\lambda\sigma)\leqslant\frac{4}{9}\frac{(1+s^2)}{(\lambda-s)^2},\qquad\lambda>s.$$

3.6　ベルジュの不等式　X_1,X_2 を一対の確率変数とし $E(X_i)=\mu_i$，$V(X_i)=\sigma_i^2$，$\cos(X_1,X_2)=\rho\sigma_1\sigma_2$ とする．このとき

$$P\left\{\max\left(\frac{|X_1-\mu_1|}{\sigma_1},\ \frac{|X_2-\mu_2|}{\sigma_2}\right)\geqslant\lambda\right\}\leqslant\frac{1+\sqrt{1-\rho^2}}{\lambda^2}.$$

3.7　$x>0$ について

$$\int_x^\infty e^{-y^2/2}dy\leqslant\frac{1}{x}e^{-x^2/2}=\frac{1}{x}\int_x^\infty ye^{-y^2/2}dy$$

であることを証明せよ．

3.8　X_1,X_2 は独立で X_i の範囲は $[0,1]$ とする．このとき $\varepsilon>0$ について

$$P\left\{\frac{1}{n}\sum[X_i-E(X_i)]\geqslant\varepsilon\right.\leqslant e^{-2n\varepsilon^2}.$$

4　極限の問題

4.1　シェッフェの収束定理［(xv), 2c. 4］は極限関数 $p(x)$ が確率密度関数でないときは必ずしも成り立たないことを例で示せ．

4.2　$X_n\xrightarrow{L}X$ とし，μ_n を X_n の中央値とする．このとき μ_n の極限の値は X の中央値である．

4.3　ヒンチンの定理（大数の弱法則）　$X_1,X_2,\cdots$ を互いに独立で同一の分布に従う確率変数列とする．このとき $E(X_i)=\mu$ が有限ならば $\bar{X}_n\xrightarrow{P}\mu$ である．

［ヒント：$\phi(t)$ を X_i の特性関数とする．このとき $\bar{X}_n$ の特性関数は $[\phi(t/n)]^n$ である．$E(X)$ は存在するのだから

$$\phi\left(\frac{t}{n}\right)=1+\frac{\mu it}{n}+o\left(\frac{t}{n}\right)$$

であり

$$\lim_{n\to\infty}\left[1+\frac{\mu it}{n}+o\left(\frac{1}{n}\right)\right]^n=e^{i\mu t}$$

であって，これは μ において退化している分布の特性関数である．］

4.4　コルモゴロフの定理（1929）　$X_1,X_2,X_3,\cdots$ を互いに独立で同一の分布 $F(x)$ に従う確率変数とし $\bar{X}_n=(X_1+\cdots+X_n)/n$ とする．このとき定数列 $\mu_1,\mu_2,\cdots$ が存在して $\bar{X}_n-\mu_n\xrightarrow{P}0$ であるための必要十分条件は

$$\int_{|x|>t}dF(x)=o(1/t),\qquad t\to\infty$$

である．この条件がみたされるならば，つねに $\mu_n=\int_{-n}^{n}xdF(x)$ ととれる．もしさらに一般化した平均

$$\mu=\lim_{t\to\infty}\int_{-t}^{t}xdF(x)$$

が存在するならば，$\bar{X}_n\xrightarrow{P}\mu$ である．コーシー分布の場合には必要十分条件がみたされないことを示せ．したがって，一般化した平均は存在するが，$\bar{X}_n$ はどんな定数にも確率収束することはない．

4.5　$X_1,X_2,\cdots$ を独立な確率変数とし $E(X_i)=\mu_i$ とする．$\bar{X}_n=(X_1+\cdots+X_n)/n$

とし $\bar{\mu}_n = (\mu_1 + \cdots + \mu_n)/n$ とする．このとき $\bar{X}_n - \bar{\mu}_n \xrightarrow{P} 0$ は，正の $\delta \leqslant 1$ について

$$\frac{1}{n^{1+\delta}} \sum_{k=1}^{n} |X_k|^{1+\delta} \to 0, \qquad n \to \infty$$

のときに成り立つことを示せ．

4.6　$X_1, X_2, \cdots$ を独立な確率変数とし，$E(X_i) = \mu_i$，$V(X_i) = \sigma_i^2$ とする．いま $S_n = X_1 + \cdots + X_n$，$M_n = \mu_1 + \cdots + \mu_n$，$V_n^2 = \sigma_1^2 + \cdots + \sigma_n^2$ とするとき，$(S_n - M_n)/V_n$ の分布関数が $N(0,1)$ に収束するための十分条件は，正の δ について

$$V_n^{-(2+\delta)} \sum_{k=1}^{n} |X_k|^{2+\delta} \to 0, \qquad n \to \infty$$

となることである．［ヒント：この十分条件がリンドバーグ＝フェラーの条件［(iii), 2c.5］を意味することを示せばよい．］

4.7　**多変量中心極限定理**　$\mathbf{X}_1, \mathbf{X}_2, \cdots$ を互いに独立な k 次元の確率変数列とし $E(\mathbf{X}_i) = 0$，$D(\mathbf{X}_i) = \Sigma_i$ とする．$n \to \infty$ のとき

$$\frac{1}{n} \sum \Sigma_i \to \Sigma \neq 0$$

であり，$\varepsilon > 0$ について

$$\frac{1}{n} \sum_{i=1}^{n} \int_{\|\mathbf{X}\| > \varepsilon\sqrt{n}} \|\mathbf{X}\|^2 dF_i \to 0$$

であるとする．ただし，F_i は $\mathbf{X}_i$ の分布関数で，$\|\mathbf{X}\|$ はベクトル $\mathbf{X}$ のユークリッド・ノルムとする．このとき確率変数 $(\mathbf{X}_1 + \cdots + \mathbf{X}_n)/\sqrt{n}$ は平均0，分散共分散行列 Σ の k 変数正規分布に収束する．

4.8　$X_1, X_2, \cdots$ を互いに独立でそれぞれ分布関数 $F_1, F_2, \cdots$ をもつ確率変数とする．$\bar{X}_n \xrightarrow{P} \mu$ となる必要十分条件は

(a)　$\sum \int_{|x|>n} dF_k \to 0$

(b)　$\frac{1}{n} \sum \int_{|x|<n} x^2 dF_k \to 0$

(c)　$\frac{1}{n^2} \sum \left\{ \int_{|x|<n} x^2 dF_k - \left(\int_{|x|<n} x dF_k \right)^2 \right\} \to 0$

である．これから期待値の存在が必要でないことがわかる．証明は Loève(1955) を見よ．

4.9　確率変数 $X, X_1, X_2, \cdots$ が一様に有界なものとする．このとき $X_n \xrightarrow{P} X$ の必要十分条件は $n \to \infty$ で $\lim E(X_n - X)^2 = 0$ であることを示せ．

4.10　すべての実数 t について

$$\lim_{n\to\infty} \int e^{tx} dF_n(x) = \int e^{tx} dF(x) = M(t)$$

であり，$M(t)$ が整関数（すなわち，複素変数の整関数 $M(z)$ が存在し，実数軸上で $M(t)$ に等しくなる）ならば，$F_n \to F$ である．証明は Kac (1959) を見よ．

4.11　$X_n \xrightarrow{L} X$ とし，X と Y とが退化しない確率変数であるとき $a_n > 0$，b_n が存在して $a_n X_n + b_n \xrightarrow{L} Y$ であるとする．このとき，定数 a と b とが存在して $aX + b \overset{L}{=} Y$（$\overset{L}{=}$ は分布が同一であることを示す）．

4.12　**不変性原理**　互いに独立で同一の分布に従う確率変数の列 $Y_1, Y_2, \cdots$ があって $E(Y_1) = 0$，$0 < E(Y_1^2) = \sigma^2 < \infty$ であり

$$\frac{Y_1 + \cdots + Y_n}{\sqrt{n}} \xrightarrow{L} Y$$

であるならば，互いに独立で同一の分布に従い $E(X_1) = 0$，$0 < E(X_1^2) < \infty$ であるような確率変数列についても同一の極限分布が対応する．

4.13　$X_n \xrightarrow{P} X$ であるならば部分列 X_{n_k} が存在し $X_{n_k} \to X$ であることを示せ．

4.14　互いに独立な確率変数 $X_1, X_2, \cdots$ について

$$\sum_1^n X_i \xrightarrow{P} \Leftrightarrow \sum_1^n X_i \xrightarrow{\text{a.s.}}$$

である．

4.15　$X_1, X_2, \cdots$ が互いに独立で $E(X_k) = 0$，$V(X_k) < \infty$ とする．さらに定数 $b_1 < b_2 < \cdots < \cdots \to \infty$ が存在し，$\sum_1^\infty [V(X_k)/b_k^2] < \infty$ としよう．このとき $\sum b_n^{-1} X_n$ は収束し

$$\frac{X_1 + \cdots + X_n}{b_n} \xrightarrow{\text{a.s.}} 0$$

である．［ヒント：クロネッカーの補助定理を使う．x_k を任意の数列とし $0 < b_1 < b_2 < \cdots \to \infty$ とする．このときもし $\sum b_k x_k$ が収束するならば $\{(x_1 + \cdots + x_n)/b_n\} \to 0$ である．］

4.16　$E|X_n - X|^r \to 0$ ならば $X_n \xrightarrow{\text{r.m.}} X$（$r$ 次の期待値が収束する）と書く．次式を証明せよ．

(a)　$X_n \xrightarrow{\text{r.m.}} X \Rightarrow E|X_n|^r \to E|X|^r$,

(b)　$X_n \xrightarrow{\text{r.m.}} X \Rightarrow X_n \xrightarrow{\text{s.m.}} X$,　$0 \leqslant s \leqslant r$,

(c)　$X_n \xrightarrow{\text{r.m.}} X \Rightarrow X_n \xrightarrow{P} X$.

4.17　$X_n \xrightarrow{P} 0$ であるが $E(X_n) \to c \neq 0$ である例を示せ．

5　**0-1 法則**　$X_1, X_2, \cdots$ を確率変数とし，$g(X_1, X_2, \cdots)$ をそれらの変数のベール関数とし，条件つき確率

$$P(g = 0 | X_1 = x_1, \cdots, X_n = x_n)$$

がすべての n およびすべてのベクトル $(x_1, \cdots, x_n)$ について条件なしの確率 $P(g = 0)$ に等しいとする．このとき $P(g = 0)$ は0または1である．

特に $X_1, X_2, \cdots$ が互いに独立ならば，数列 $\sum_1^\infty X_n$ の収束の確率は0または1である．

6　$F_1(x)$ と $F_2(x)$ が有界変動の単調関数で，c_1, c_2 が任意の定数であるとき

$$\int_a^b f(x) d[c_1 F_1(x) + c_2 F_2(x)] = c_1 \int_a^b f(x) dF_1(x) + c_2 \int_a^b f(x) dF_2(x)$$

である．

7　X_1, X_2 が互いに独立な分布 F_1, F_2 に従うなら，$Z = X_1 + X_2$ および $Y = X_1/X_2$ の分布関数

$$F(z) = \int F_1(z - x) dF_2(x) = \int F_2(z - x) dF_1(x)$$

$$F(y)=\int_0^\infty F_1(yx)dF_2(x)+\int_{-\infty}^0[1-F_1(yx)]dF_2(x)$$

である．

8　1e の不等式を用いて確率変数 X, Y に関し次のことを示せ．

（a）$E|X+Y|^r \leqslant C_r(E|X|^r+E|Y|^r)$

ここで $C_r=1$ $(r\leqslant 1)$，$=2^{r-1}$ $(r>1)$ とする．

（b）ヘルダーの不等式

（i）$E|XY|<(E|X^r|)^{1/r}(E|Y^s|)^{1/s}$，　$1<r<\infty$，$\dfrac{1}{r}+\dfrac{1}{s}=1.$

（ii）$(E|XY|)^2\leqslant E(X^2)E(Y^2)$，　コーシー＝シュヴァルツ．

（iii）$E|X|\leqslant(E|X|^r)^{1/r}$，　$r>1.$

（iv）$(E|X|^t)^{1/t}\leqslant(E|X|^r)^{1/r}$，　$0<t\leqslant r.$

（c）ミンコフスキの不等式

$$(E|X+Y|^r)^{1/r}\leqslant(E|X|^r)^{1/r}+(E|Y|^r)^{1/r},\qquad 1\leqslant r<\infty.$$

（d）$\log E|X|^r$ は r についての凸関数である．

（e）イェンセンの不等式　g が凸関数で $E(X)$ が存在するならば

$$g[E(X)]\leqslant E[g(X)].$$

9　無限分解可能な分布　分布関数 F が無限分解可能である（弱い意味で）とはすべての n について，退化しない分布関数 $F_1,\cdots,F_n$ のたたみこみとして書けるということである．いいかえると，分布 F をもつ確率変数 X が退化しない独立な n 個の確率変数の和としてあらわされることである．すべての n について X が n 個の互いに独立で同一の分布に従う確率変数の和として表現されるならば，分布関数 F は無限分解可能であるという．無限分解可能な分布について次のことを証明せよ．

9.1　任意の n について X の特性関数は，他の特性関数の n 乗である．

9.2　X の特性関数は，0になることは決してない．

9.3　正規分布およびポアソン分布は無限分解可能である．

9.4　互いに独立で無限分解可能な確率変数の和は無限分解可能である．

9.5　無限分解可能な分布の系列の（弱い意味の）極限はそれ自身無限分解可能である．

［注意：無限分解可能な分布の正準形表示については，Doob, Gnedenko, Loève, Ramachandran などの確率論の書物を参照されたい．］

10　離散分布

10.1　$P_1(t),\cdots,P_k(t)$ を k 個の独立な離散確率分布 $X_1,\cdots,X_k$ の確率母関数とする．$X_1+\cdots+X_k$ なる和の確率母関数は積 $P_1(t)\cdots P_k(t)$ となることを示せ．

10.2　$S_N=X_1+\cdots+X_N$ を N 個の独立な離散確率分布の和とし，それらの確率母関数はすべて同じ $P(t)$ としよう．さらに N も確率変数で確率母関数が $G(t)$ であるとする．このとき独立な確率変数の N 個（N は確率変数）の和として定義される確率変数 S の確率母関数は $G[P(t)]$ であることを示せ．

10.3　**10.2** の結果を用いて，X_i がパラメータ λ のポアソン分布に，N がパラメータ μ のポアソン分布に従う場合の S の分布を求めよ．

10.4　上と同様にして，X_i がパラメータ $(\pi,\ n=1)$ の2項分布に，N がパラメータ μ のポアソン分布に従うとき S はパラメータ $\mu\pi$ のポアソン分布に従うことを示せ．

11　以下の条件つき確率を求めよ．

11.1　X と Y とを独立な2項分布とし，パラメータをそれぞれ (π,n_1)，(π,n_2) とする．このとき，$X+Y$ を与えたときの (X,Y) の条件つき分布は超幾何分布で π とは無関係である．

11.2　X と Y を独立な負の2項分布とし，パラメータをそれぞれ (ρ,κ_1)，(ρ,κ_2) とする．このとき，$X+Y$ を与えたときの (X,Y) の条件つき確率は次に定義される負の二項分布に従う．

$$\binom{X+Y}{X}\frac{B(\kappa_1+X,\kappa_2+Y)}{B(\kappa_1,\kappa_2)}$$

ただし $B(m,n)$ はベータ関数である．

12　代表的な分布の特性関数

	密度関数	特性関数
正規：	$\dfrac{1}{\sqrt{2\pi}\,\sigma}e^{-(x-\mu)^2/2\sigma^2}$	$e^{i\mu t-\sigma^2t^2/2}$
ラプラス：	$\dfrac{1}{2\sigma}e^{-\lvert x-\mu\rvert/\sigma}$	$e^{it\mu}(1+\sigma^2t^2)^{-1}$
長方形：	$\dfrac{1}{\theta}$，$a<x<a+\theta$	$e^{ita}\dfrac{e^{it\theta}-1}{it\theta}$
ガンマ：	$\dfrac{\alpha^p}{\Gamma(p)}x^{p-1}e^{-\alpha x}$	$\left(\dfrac{\alpha}{\alpha-it}\right)^p$
ベータ：	$\dfrac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)}x^{p-1}(1-x)^{q-1}$	$\dfrac{\Gamma(p+q)}{\Gamma(p)}\displaystyle\sum_{j=0}^{\infty}\dfrac{(it)^j\Gamma(p+j)}{\Gamma(p+q+j)\Gamma(j+1)}$
コーシー：	$\dfrac{\delta}{\pi}\dfrac{1}{\delta^2+(x-\mu)^2}$	$e^{it\mu}e^{-\lvert t\rvert\delta}$
3角形 $\lvert x\rvert<a$：	$\dfrac{1}{a}\left(1-\dfrac{\lvert x\rvert}{a}\right)$	$2\dfrac{1-\cos at}{a^2t^2}$
3角形 $-\infty<x<\infty$：	$\dfrac{1}{\pi}\dfrac{1-\cos ax}{ax^2}$	$\begin{cases}1-\lvert t\rvert/a & \lvert t\rvert\leqslant a\text{ のとき}\\ 0 & \lvert t\rvert>a\text{ のとき}\end{cases}$
双曲余弦 $-\infty<x<\infty$：	$\dfrac{1}{\pi\cosh x}$	$\dfrac{1}{\cosh(\pi t/2)}$
長方形 $\lvert x\rvert<\theta$：	$\dfrac{1}{2\theta}$	$\dfrac{\sin\theta t}{\theta t}$
退化：	$X=a$（確率1で）	e^{ita}

13　大文字，小文字の O の概念　$\{r_n\}$ を正の実数列とし，$\{X_n\}$ を確率変数列とする．$X_n=O_p(r_n)$ もしくは X_n/r_n が確率的に有界であるというのは各 ε にたいして M_ε および N_ε が存在し $N>N_\varepsilon$ について $P[|X_n|/r_n>M_\varepsilon]<\varepsilon$ となることをいう．また，$X_n/r_n\xrightarrow{P}0$ ならば $X_n=o_p(r_n)$ と書く．

$n\to\infty$ のとき $r_n\to 0$ であるような r_n について $X_n-a=o_p(r_n)$ とし，$g(x)$ は a で k 次のテイラー展開可能，すなわち

$$g(x)=T_k(x,a)+o(|x-a|^k)$$

とする．このとき $g(X_n) - T_k(X_n, a) = o_p(r_n^k)$ であることを示せ．

14　Y を単位区間 $(0, 1)$ で一様に分布する確率変数とする．Y を十進展開したときの逐次の数字を $X_1, X_2, \cdots$ とする．すなわち

$$Y = \frac{X_1}{10} + \frac{X_2}{10^2} + \cdots + \frac{X_n}{10^n} + \cdots$$

このとき，$X_1, X_2, \cdots$ は0から9までの整数値を一様にとる独立な離散確率変数であることを示せ．したがって，任意の整数 k について，Y を十進展開したときの k の相対頻度が 1/10 であるような数値 Y の集合は確率1をもつという結論を導け．このことは，1/3 を十進展開すると，3しか出現しないことと矛盾するか？

15　$\{a_n\}$ を $n \to \infty$ のとき $a_n \to a$，$f(a_n) \to b$ なる数列とする．いま $X_n \xrightarrow{P} a$ となるような確率変数の系列 $\{X_n\}$ を考える．このとき，$f(X_n) \xrightarrow{P} b$ を示せ．また $X_n/n \xrightarrow{P} 0$ ならば，$\max(X_1/n, \cdots, X_n/n) \xrightarrow{P} 0$ を意味するか？［この結果は確率的でない数列にたいしては正しい．］

16　次の例を示せ．

（a）　相異なる特性関数が有限区間の上で一致する場合．

（b）　$F * F_1 = F * F_2$ であり $F_1 = F_2$ でない場合．

17　**積率の問題**　次の2つの正の確率変数 X の積率が一致することを示せ．

$$p_1(x) = (2\pi)^{-1/2} x^{-1} \exp[-(\log x)^2/2]$$
$$p_2(x) = p_1(x)[1 + a \sin(2\pi \log x)], \qquad -1 \leqslant a \leqslant 1.$$

この興味ある例は，積率から確率分布が一意的に定まらないことを示すもので，C. C. ヘイドによる．

18　X_1, X_2 が互いに独立で同一の分布に従うとき，$t > 0$ について

$$P\{|X_1 - X_2| > t\} \leqslant 2P\{|X_1| > t/2\}$$

である．$X_1, \cdots, X_n$ が独立で同一の分布 F に従うならば

$$2P\{|X_1 + \cdots + X_n| \geqslant t\} \geqslant 1 - \exp\{-n[1 - F(t) + F(-t)]\}$$

である．

19　X, Y, Z が確率変数のとき，**2b. 3** の方法を用いて以下の条件つき期待値に関する結果を証明せよ．

（a）　$\mathrm{Cov}(X, Y) = \mathrm{Cov}[X, E(Y|X)]$

（b）　$E(Z) = E\{E[E(Z|Y)|X]\}$,

$V(Z) = E\{E[V(Z|Y)|X]\} + E\{V[E(Z|Y)|X]\} + V\{E[E(Z|Y)|X]\}$,

$\mathrm{Cov}(X, Z) = \mathrm{Cov}\{X, E[E(Z|Y)|X]\}$.

20　$(C_1, \cdots, C_k)$ を標本空間の分割とするとき

$$P(A|B) = \sum_1^k P(AC_i|B) = \sum_1^k P(C_i|B)P(A|BC_i)$$

を示せ．

21　確率空間 $(\Omega, \mathscr{B}, P)$ を考える．Y を点 y の空間，$\mathscr{A}$ を Y の集合の σ 集合体，$T: \Omega \to Y$ を $A \in \mathscr{A} \Rightarrow T^{-1}A \in \mathscr{B}$ であるような関数とする．$f(\omega, y)$ を $\Omega \times Y$ の上の実数値の $\mathscr{B} \times \mathscr{A}$ 可測な関数とする．このとき

$$E[f(\omega, T(\omega))|T(\omega) = y] = E[f(\omega, y)|T(\omega) = y]$$

であることを示せ（Bahadur and Bickel (1968) を参照）．

22　$\mathbf{X}$ をベクトルパラメータ θ をもつベクトル値の確率変数とする．また，$\mathbf{T}$ を統計量とし，$\mathbf{T}$ の分布は θ の関数である ϕ にのみ関係するとする．このとき $\mathbf{T}$ が ϕ の推測に関して十分であるというのは，$\mathbf{T}$ が与えられたときの $\mathbf{X}$ の条件つき分布が，ϕ とは無関係な θ の関数のみに依存するときにいう（Rao (1965 b) 参照）．

$X_1, \cdots, X_n$ を $N(\mu, \sigma^2)$ からの n 個の独立な観測値とする．$\sum(X_i - \bar{X})^2$ はパラメータ σ^2 推測に関して十分であることを示せ．

23　$E(X_n) \to 0$，$V(X_n) \to 0$ ならば $X_n \xrightarrow{P} 0$ であることの簡単な証明を考えよ．

参考文献

Bahadur, R. R. (1954), Sufficiency and statistical decision functions, *Ann. Math. Statist.* **25**, 423–462.

Bahadur, R. R. (1955), Statistics and Subfields, *Ann. Math. Statist.* **26**, 490–497.

Bahadur, R. R. and E. L. Lehmann (1955), Two comments on "Sufficiency and statistical decision functions," *Ann. Math. Statist.* **26**, 139–142.

Bahadur, R. R. and P. J. Bickel (1968), Substitution in conditional expectation, *Ann. Math. Statist.* **39**, 377–378.

Cramer, H. (1946), *Mathematical Methods of Statistics*, Princeton University Press, Princeton, N. J.

Cramer, H. and H. Wold (1936), Some theorems on distribution functions, *J. London Math. Soc.* **11**, 290–294.

Dharmadhikari, S. W. (1965), An example in the problem of moments, *Am. Math. Monthly*, **72**, 302–303.

Doob, J. L. (1953), *Stochastic Processes*, John Wiley, New York.

Feller, W. (1966), *An Introduction to Probability Theory and Its Applications*, John Wiley, New York.

Ferguson, T. S. (1967), *Mathematical Statistics. A Decision Theoretic Approach*, Academic Press.

Gnedenko, B. V. (1962), *Theory of Probability*, Chelsea, New York.

Gnedenko, B. V. and A. N. Kolmogorov (1954), *Limit Distributions for Sums of Independent Random Variables*, Addison-Wesley, Reading, Mass.

Grenander, V. (1960), *Probabilities on Algebraic Structures*, John Wiley, New York.

Hajek, J. and A. Renyi(1950), Generalisation of an inequality of Kolmogorov, *Acta. Math. Acad. Sci. Hung.*, **6**, 281–283.

Halmos, P. R. (1950), *Measure Theory*, D. Van Nostrand, New York.

Halmos, P. R. and L. J. Savage (1949), Application of the Radon-Nikodym theorem to the theory of sufficient statistics, *Ann. Math. Statist.* **20**, 225–241.

Hamburger, H. (1920, 1921), Über eine Erweiterung des Stieltjesschen Momentenproblems, *Mathemat. Annalen* **81**, 235–319, **82**, 120–167, 167–187.

Kac, M. (1959), *Probability and Related Topics in Physical Sciences*, Interscience, London and New York.

Kolmogorov, A. N. (1929), Bemerkungen zu meiner Arbeit "Über die Summen zufälliger Grössen", *Math. Annalen* **102**, 484–488.

Kolmogorov, A. N. (1950), *Foundations of the Theory of Probability* (German edition, 1933), Chelsea, New York.

Lévy, P. (1937), *Theorie de l'addition des variables aleatoires*, Gauthier-Villars, Paris.
Linnik, Yu. V. (1960), *Decomposition of Probability Laws*. Leningrad University.
Loève. M. (1955), *Probabitity Theory*, D. Van Nostrand, Princeton, N.J.
Mann, H. B. and A. Wald (1943), On stochastic limit and order relationships, *Ann. Math. Statist.* **14**, 217–226.
Munroe, M. E. (1959), *Introduction to Measure and Integration*, Addison-Wesley, Reading, Mass.
Okamoto Masashi (1959), A convergence theorem for discrete probability distributions, *Inst. Stat. Math Tokyo* **11**, 107–112.
Parthasarathy, K. R. (1967), *Probability Measures on Metric Spaces*, Academic Press.
Pathak, P. K. (1966), Alternative proofs of some theorems on characteristic functions, *Sankhyā A*, **28**, 309–315.
Pratt, J. W. (1960), Extension of dominated convergence theorem, *Ann. Math. Statist.*, **31**, 74–77.
Ramachandran, B. (1967), *Advanced Characteristic Functions*, Statistical Publishing Society.
Savage, L. J. (1970), An unpublished note.
Scheffe, H. (1947), A useful convergence theorem for probability distributions, *Ann. Math. Statist.* **18**, 434–458.
Sudakov, V. N. (1971), A note on marginal sufficiency, *Doklady*, 198.
Varadarajan, V. S. (1958), A useful convergence theorem, *Sankhyā*, **20**, 221–222.
Wald, A. and J. Wolfowitz (1944), Statistical tests based on permutations of the observations, *Ann. Math. Statist.* **15**, 358–372.

第3章　連続型確率模型

まえがき　第2章で分布関数によって規定される確率変数の概念を導入した．実際には分布関数 F は未知であるから，確率変数の具体的な実現値をもとに F について推測することが問題となる．しかし，C を分布関数の特定なクラスとするとき，先験的に $F \in C$ であることが知られているならば大変助かるであろう．なぜなら，そのような場合，われわれの調査の対象はすべての分布関数の広大なクラスではなく特定のクラスにかぎられるからである．ここで用いる $F \in C$ という情報を規定と呼ぶ．ときには，観測値を生成する機構に関する知識からそこに含まれる無作為性を適当に特徴づけをることによって，そのクラス C をきめることができる．この特徴づけをモデルと呼ぶ．

たとえば n 個の壺があり，それぞれ未知であるが一定の数の白球と黒球を含んでいるとしよう．各壺から1つの球を取り出し，その色を記録する．取り出した n 個のうち r 個が白である確率はいくらであろうか？　どの壺の中の球もすべて区別できるとし，各壺には a 個の白球と b 個の黒球が入っているとしよう．すると n 個の球について $(a+b)^n$ 通りの可能な組合せができる．もし十分にかき混ぜた後，各壺から1つずつ球をとり出すと，$(a+b)^n$ 通りのすべての組合せが同等に起こるであろうという単純な仮定をすることができる．すると，r 個の白球をもつ組合せの起こる割合は

$$\binom{n}{r}\pi^r(1-\pi)^{n-r}$$

であることが容易にわかる．ただし $\pi = a/(a+b)$ は未知の割合である．このようにして，簡単なモデルから2項分布 (2a. 5. 3) を導くことができる．

本章では，よく知られた連続分布を導く数多くのモデルを考える．またこれらの分布の性質も調べる．さらに，観測値の関数で統計的推測に用いられるものの確率分布も求める．

1変量および多変量の解析におけるすべての分布の問題の解決の手がかりを与える 3b. 5 の最小2乗法の3つの基本定理をくわしく学ぶことは，読者にとって有益であろう．本章の終わりに問題として与えたこれらの定理のいくつかの拡張もまた有用である．

3d で2変量正規分布を導入する．そこでは，いくつかの重要な統計量の標本分布を変数変換によって求めている．この方法を多変量の場合に拡張することもできるが，しかしこの拡張は行わない．それは，多変量解析に関する第8章で示すようにもっと巧みな方法が利用できるからである．

変数変換　Xを密度関数が p_X である確率変数とする．$Y=f(X)$の密度関数は何か？定義より

$$P(Y<y_0)=P\{x:f(x)<y_0\}=\int_A p_X(x)dx$$

である．ここで A は $f(x)<y_0$ となる領域である．次の仮定をおこう．（a）変換 $y=f(x)$ が1対1で逆変換 $x=g(y)$ が存在する．（b）g の導関数 $g'(y)$ が存在し，連続である．するとリーマン積分での変数変換法により

$$\int_A p_X(x)dx=\int_{y<y_0} p_X[g(y)]|g'(y)|dy \tag{3.1}$$

となる．したがって定義から Y の密度関数は $p_X[g(y)]|g'(y)|$ となる．

変換は1対1でないが，x の範囲を有限個の互いに排反な領域 $R_1,\cdots,R_k$ に分割し，Y の範囲でそれぞれに対応する領域を $R'_1,\cdots,R'_k$（重なっていてもよい）とするとき，R_i から R'_i への変換は1対1となるようにできるとしよう．$x=g_i(y)$ を R'_i から R_i への変換とし，$g'_i(y)$ は連続とする．このとき y での確率密度は

$$\sum_{i=1}^{k}\pi_i p_X[g_i(y)]|g'_i(y)| \tag{3.2}$$

となる．ここで $y\in R'_i$ ならば $\pi_i=1$，その他の場合は $\pi_i=0$ である．

たとえばもし $Y=X^2$ ならば次の2つの領域が考えられる．$R_1=(-\infty,0)$，$R_2=(0,\infty)$．Y の空間で対応する領域は $R'_1=(0,\infty)$，$R'_2=(0,\infty)$ でありこれらは完全に重なっている．(R_1,R'_1) に対応する逆関数は $x=-\sqrt{y}$ で，(R_2,R'_2) にたいしては $x=\sqrt{y}$ である．したがって任意の $y\in R'_1,R'_2$ での確率密度は

$$[p_X(\sqrt{y})+p_X(-\sqrt{y})]\left(\frac{1}{2}\cdot y^{-\frac{1}{2}}\right)$$

となる．

この結果は多変数の場合に同様の方法で拡張できる．$Y_i=f_i(X_1,\cdots,X_n)$，$i=1,\cdots,n$ を与えられた変換，$p(x_1,\cdots,x_n)$ を $X_1,\cdots,X_n$ の密度関数とする．R_i が X の空間にあり，R'_i が Y の空間にあり，各対 (R_i,R'_i) に1対1の対応がつく領域 $(R_1,R'_1),\cdots,(R_k,R'_k)$ を考えよう．

$$x_r=g_{ri}(y_1,\cdots,y_n),\qquad r=1,\cdots,n$$

を (R_i,R'_i) の逆変換とする．g_{ri} の偏導関数が連続で，変換のヤコビアン (J_i) が R'_i で零でなければ，$y_1,\cdots,y_n$ での確率密度は

$$\sum \pi_i p(g_{1i},\cdots,g_{ni})|J_i| \tag{3.3}$$

である．ここで $(y_1,\cdots,y_n)\in R'_i$ ならば $\pi_i=1$，その他の場合は $\pi_i=0$ である．また $|J_i|$ はヤコビアンの絶対値である．

記号　第2章と前述の変数変換の例では，確率変数をXであらわし，その密度関数を $p_X(x)$，あるいは背後の確率変数について混乱の恐れがないときは単に $p(x)$ であらわした．このことは厳密には守られていない．というのは統計学では R,r,F,f,T,t などの大文字も小文字も確率変数の特定の関数をあらわすのに使われているからである．それゆえ，確率変数をあらわすのに任意の記号を自由に用いることにしよう．これを許すと，密度関数のなかの記号は形式的なものであるから，確率変数とその密度関数の表現のなかでは同じ記号を用いることにすると便利である．そうすると，x は密度関数 $f(x)$ をもつ確率変数であるということができ，x^* のような異なった記号を用いてその密度関数が $f(x^*)$ であると言う必要はなくなる．このような協定は，密度関数にたいして与えた表現が，いま考えている特定の確率変数（きめられた記号をもつ）の密度関数であるとわかる場合には都合がよい．さらに以下で考える大部分の命題では，確率変数の関数の密度関数をきめるために式 (3.1), (3.2), (3.3) を適用している．この目的にたいしては変換をあらわすときも，ヤコビアンの評価をするときも，どのような記号（確率変数に使われるのと同じものあるいは確率変数Xにたいしてxのような対応する記号）を用いるかは重要でない．しかしながら，最初の前提から式 (3.1) ないし (3.3) を導く過程においては，このような区別は事柄を明確にさせるために必要である．

われわれは分布の族を引用するときしばしば，すべての正規分布のクラスにはNのような記号を，平均μ，分散σ^2もつ正規分布には $N(\mu,\sigma^2)$ のように，その母数を追加指定する記号を用いる．X が $N(\mu,\sigma^2)$ の分布に従う確率変数であることをあらわすのに $X\sim N(\mu,\sigma^2)$ という記号を用いる．$N(\mu,\sigma^2)$ の密度関数は，$N(x|\mu,\sigma^2)$ あるいは X 自身を使って $N(X|\mu,\sigma^2)$ と書く．

3a　1変量模型

3a.1　正規分布

（i）**標的の射撃（ハーシェルの仮説）**　標的（理想的には1点）に向かって発した散弾の分布を考えよう．(X,Y) を2つの直交軸に関して，標的からの散弾のずれ（誤差）をあらわす（確率変数の）座標とする．

次の仮定が成り立つとしよう：

（a）誤差XとYの周辺密度関数 $p(x),q(y)$ は連続である．

（b）(x,y) での確率密度は原点からの距離 $r=(x^2+y^2)^{1/2}$ にのみ依存する（半径に関する対称性）．

（c）　x 方向と y 方向の誤差は独立である．

このとき任意の方向のずれ Z の密度関数は正規密度関数

$$\frac{1}{\sigma\sqrt{2\pi}}e^{-z^2/2\sigma^2}$$

である．

(b) と (c) を用いると (x, y) での確率密度は

$$p(x)q(y) = s(r), \qquad r^2 = x^2 + y^2 \tag{3a.1.1}$$

となる．$x = 0$ とおけば関数 s と q は比例し，$y = 0$ とおけば s と p は比例することがわかる．したがって関数方程式 (3a.1.1) は $f(x) = \log[p(x)/p(0)]$ と書けば

$$f(x) + f(y) = f(r), \qquad r^2 = x^2 + y^2 \tag{3a.1.2}$$

となる．さらに (3a.1.2) で $y = 0$，$x = -x$ とおくと $f(x) = f(-x) = f(|x|)$ が得られる．これより $x^2 = x_1^2 + x_2^2$ ならば

$$f(r) = f(y) + f(x_1) + f(x_2), \qquad r^2 = y^2 + x_1^2 + x_2^2$$

であり，一般的には

$$f(r) = f(x_1) + \cdots + f(x_k), \qquad \sum x_i^2 = r^2$$

となる．$k = n^2$，$x = x_1 = \cdots = x_k$ とすると

$$f(nx) = n^2 f(x)$$

すなわち $x = 1$ のとき $f(n) = n^2 f(1)$ となる．m が整数のとき $x = m/n$ にたいし

$$n^2 f\left(\frac{m}{n}\right) = f\left(n\frac{m}{n}\right) = f(m) = m^2 f(1)$$

すなわち $c = f(1)$ とすると

$$f\left(\frac{m}{n}\right) = c\left(\frac{m}{n}\right)^2$$

となる．したがってすべての有理数 x について $f(x) = cx^2$ となり，連続性によってすべての x にたいし，この関係が成り立つ．これより

$$p(x) = p(0)e^{cx^2} \tag{3a.1.3}$$

となる．(3a.1.3) が密度関数となるためには c は負でなければならないので $-1/2\sigma^2$ と書ける．(3a.1.3) を $-\infty$ から ∞ まで積分し，その結果が 1 になるようにすると $p(0) = 1/\sigma\sqrt{2\pi}$ となり

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}}e^{-x^2/2\sigma^2} \tag{3a.1.4}$$

となる．これがすでに **2c.5** で導入した $E(x) = 0$，$V(x) = \sigma^2$ をもつ有名な正規分布，$N(0, \sigma^2)$ である．

誤差 X, Y の同時密度関数は

$$p(x)p(y) = \frac{1}{\sigma^2 2\pi}e^{-(x^2+y^2)/2\sigma^2} \tag{3a.1.5}$$

である．任意の方向 $(\cos\theta, \sin\theta)$ での誤差は $Z = X\cos\theta + Y\sin\theta$ である．Z の密度関数を求めるために変換

$$z = x\cos\theta + y\sin\theta, \qquad u = x\sin\theta - y\cos\theta$$

を考えよう．変換のヤコビアンは $D(z, u)/D(x, y) = -1$ である．密度関数 (3a.1.5) は

$$\frac{1}{\sigma^2 2\pi}e^{-(z^2+u^2)/2\sigma^2}$$

に変換される．これは U と Z が独立であることを示している．Z の密度関数は

$$\frac{1}{\sigma\sqrt{2\pi}}e^{-z^2/2\sigma^2}$$

となり，求める結果が得られた．

$E(X) = \mu$，分散 σ^2 の正規分布は特性関数 (2b.4.11) が

$$e^{it\mu-\sigma^2t^2/2} \tag{3a.1.6}$$

であり密度関数は

$$N(x|\mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}}e^{-(x-\mu)^2/2\sigma^2}$$

である．

（ii）　**マックスウェルの仮説**　マックスウェルは次の仮定の下で分子の速度分布を導くときに正規分布を得た．

（a）　3つの直交する方向の速度の成分 u, v, w は独立に分布する．

（b）　u, v, w の周辺分布は同一である．

（c）　相空間は等方性である．つまり与えられた速度成分をもつ分子の密度は，全速度の関数であって，その方向にはよらない．

$f(\cdot)$ が速度の任意の成分の密度関数をあらわすとすると (3a.1.1) と同様に (a) から (c) の仮定により関数方程式

$$f(u)f(v)f(w) = g(V), \qquad V^2 = u^2 + v^2 + w^2 \tag{3a.1.7}$$

が導かれる．したがって $f(u)$ は (3a.1.4) の形をとり，これは任意の1つの速度成分にたいして正規分布であるから

$$g(V) = \text{const}\cdot e^{-\alpha(u^2+v^2+w^2)} \tag{3a.1.8}$$

となる．

仮定 (a), (b) は正規分布を導くのに必要でないことが J. W. Mayer and M. G. Mayer (1940) によって指摘されている．

仮定 (c) を用いて $g(V)$ を点 (u, v, w) での確率密度とする．物理学の法則(分子の衝突時の運動エネルギーの保存則) から

$$g(V_1)g(V_2) = g(V_1')g(V_2') \tag{3a.1.9}$$

が導かれる．ここで速度 V_1, V_2, V_1', V_2' は

$$V_1^2 + V_2^2 = (V_1')^2 + (V_2')^2 \tag{3a.1.10}$$

の関係を満足する．制約 (3a.1.10) があるとき，条件 (3a.1.9) の下での唯一の解は (3a.1.8) である．つまり速度の各成分は正規分布に従い，各成分の分布は互いに独立である．

（iii）　**2項分布の極限（ドモアブルの定理）**　成功の確率が π である n 回のベルヌイ試行で r を成功の数とし，確率変数を

$$x = \frac{r - n\pi}{\sqrt{n\pi\phi}}, \qquad \phi = 1 - \pi$$

と定義する．このとき x の極限分布は正規分布になる．

(2b.4.8) で示したように r の特性関数は $E(e^{itr}) = (\phi + \pi e^{it})^n$ であり，x の特性関数は

$$\begin{aligned} E(e^{itx}) &= E(e^{it(r-n\pi)/\sqrt{n\pi\phi}}) = e^{-itn\pi/\sqrt{n\pi\phi}} E(e^{itr/\sqrt{n\pi\phi}}) \\ &= e^{-itn\pi/\sqrt{n\pi\phi}} (\phi + \pi e^{it/\sqrt{n\pi\phi}})^n = (\phi e^{-it\pi/\sqrt{n\pi\phi}} + \pi e^{it\phi/\sqrt{n\pi\phi}})^n \end{aligned} \tag{3a.1.11}$$

である．$e^{iz} = 1 + (iz) + \cdots + (iz)^k/k! + o(z^k)$ に注意し，(3a.1.11) のなかの指数を展開すると

$$f_x(t) = \left[1 - \frac{t^2}{2n} + o\left(\frac{t^2}{n}\right)\right]^n \to e^{-t^2/2}, \qquad n \to \infty \text{ のとき} \tag{3a.1.12}$$

となる．この右辺は $N(0,1)$ の特性関数である．したがって連続性定理 [(iv), 2c.4] により x の極限分布は正規分布である．つまり (3a.1.12) の結果から，n が大きければ2項確率変数 r の分布関数は平均 $n\pi$，分散 $n\pi\phi$ の正規確率変数の分布関数で近似できる．

しかし (3a.1.12) の結果は **2c.5** で述べた中心極限定理の特別な場合にすぎない．それは，有限の平均と分散をもち，互いに独立に同一の分布に従う n 個の確率変数の平均値の極限分布は，ある適当に定義された意味で正規分布であると言っている．

（iv）　**ハーゲンの仮説（誤差の理論）**　ハーゲンは次の仮定の下で誤差の正規法則の証明を行った．

（a）　1つの誤差は，異なる原因で起こり，大きさはすべて等しい無数の微小誤差の和である．

（b）　誤差の異なる成分は互いに独立である．

（c）　誤差の各成分が正または負となる確率は等しい．

仮定 (c) により誤差の各成分はそれぞれ確率 $\frac{1}{2}$ で $\pm\varepsilon$ の値をとるので，平均は0，分散は ε^2 である．$x = \varepsilon_1 + \cdots + \varepsilon_n$ が n 個の独立な成分による全誤差ならば，

$$E(x) = E(\varepsilon_1) + \cdots + E(\varepsilon_n) = 0$$

$$V(x) = \sum V(\varepsilon_i) = n\varepsilon^2 = \sigma^2 \quad (\text{とおく})$$

となる．σ^2 が有限で一定となるように $n \to \infty$ かつ $\varepsilon \to 0$ としたときの x の極限分布を求めよう．ε_j の特性関数は $\frac{1}{2}(e^{it\varepsilon} + e^{-it\varepsilon})$ で $x = \varepsilon_1 + \cdots + \varepsilon_n$ の特性関数は $n \to \infty$ のとき

$$\left(1 - \frac{t^2}{2!}\varepsilon^2 + \frac{t^4}{4!}\varepsilon^4 + \cdots\right)^n = \left[1 - \frac{t^2}{2}\frac{\sigma^2}{n} + o\left(\frac{1}{n}\right)\right]^n \to e^{-t^2\sigma^2/2}$$

となる．これは $N(0, \sigma^2)$ の特性関数である．

（v）　**n 個の独立な正規確率変数の和**　$x_1, \cdots, x_n$ を n 個の独立な正規確率変数で x_j の平均を μ_j，分散を σ_j^2 とする．和 $x = x_1 + \cdots + x_n$ の特性関数は $x_1, \cdots, x_n$ の特性関数の積である．(3a.1.6) の結果を用いると

$$\prod_{j=1}^{n} e^{it\mu_j - \sigma_j^2 t^2/2} = e^{it\mu - \sigma^2 t^2/2}, \qquad \mu = \sum \mu_j, \quad \sigma^2 = \sum \sigma_j^2$$

が得られる．したがって和の分布はその平均が各平均の和に等しく，分散が各分散の和に等しいような正規分布に正確に従う．よって平均 $\bar{x} = (x_1 + \cdots + x_n)/n \sim N(\mu, \sigma^2)$ となる．ただし

$$\mu = \frac{\mu_1 + \cdots + \mu_n}{n}, \qquad \sigma^2 = \frac{\sigma_1^2 + \cdots + \sigma_n^2}{n^2}$$

である．

中心極限定理により確率変数の分布の性質にはある程度無関係に，確率変数の和の漸近的正規性が成立する．もし n 個の独立な確率変数の和が正確に正規分布に従うならば個々の確率変数の分布についてどのようなことがいえるだろうか？　n 個の独立な確率変数の和が正規分布に従うならば個々の確率変数は正規分布に従うという Cramer (1937) による有名な定理がある．

（vi）　**平均と分散が与えられたときの最大エントロピー**　密度関数が $p(x)$ である1次元の確率変数を考え，分布のエントロピーを

$$-\int p(x) \log p(x) dx \tag{3a.1.13}$$

で定義する．物理学での粒子の分布の研究におけるエントロピーの意義と役割は 3a.6 で論じる．条件

$$\left.\begin{aligned} &\int p(x)dx = 1, \quad \int xp(x)dx = \mu, \\ &\int (x-\mu)^2 p(x)dx = \sigma^2 \end{aligned}\right\} \tag{3a.1.14}$$

（ただし μ, σ^2 は固定されている）の下で (3a.1.13) が最大となるように範囲 $(-\infty, \infty)$ で定義された $p(x)$ を定めよう．

(1e. 6. 6) で任意の2つの密度関数 p, q にたいして

$$-\int p \log \frac{p}{q} dx \leqslant 0$$

であることが示されている．(3a. 1. 14) が p のかわりに q について成り立つように α, β, γ がきめられた

$$\log q = \alpha + \beta(x-\mu) + \gamma(x-\mu)^2$$

を選ぶと，p にたいする条件 (3a. 1. 14) を用いて

$$-\int p \log p\, dx \leqslant -\int p \log q\, dx = -\int p[\alpha + \beta(x-\mu) + \gamma(x-\mu)^2] dx$$
$$= -(\alpha + \gamma\sigma^2)$$

であることがわかる．これより $-(\alpha+\gamma\sigma^2)$ は $p(x)$ が条件 (3a. 1. 14) をみたすとき，エントロピーの一定の上限となる．一方この値は p を

$$\log p = \alpha + \beta(x-\mu) + \gamma(x-\mu)^2$$

すなわち

$$p = \exp[\alpha + \beta(x-\mu) + \gamma(x-\mu)^2] \qquad (3a. 1. 15)$$

となるように選んだとき得られる．われわれは (3a. 1. 14) をみたすように α, β, γ がきめられることを確かめなければならない．(3a. 1. 15) の形から

$$\alpha = -\log(\sqrt{2\pi}\sigma), \quad \beta = 0, \quad \gamma = -\frac{1}{2\sigma^2}$$

と選べば，(3a. 1. 14) をみたすことは明らかである．このように正規分布は確率変数の平均，分散が固定され，範囲が $(-\infty, \infty)$ のとき最大エントロピーをもつという性質によって特徴づけられる．読者は，

$$\int p \log \frac{p}{q} dx = 0$$

は p と q がほとんどいたるところで等しいことを意味していることから ((1e. 6. 6) 参照)，解 (3a. 1. 15) は本質的に一意であることを確かめることができる．

(vii)　正規確率変数の2乗　　$x \sim N(\mu, \sigma^2)$ とする．問題は $y = x^2$ の分布をきめることである．与えられた y にたいし変換が1対1でないから，x の2つの値 $(\pm x)$ が存在する．x と $-x$ の密度関数は ($dy = 2x dx$ に注意して)

$$(2\sigma\sqrt{2\pi y})^{-1} \exp\left[-\frac{(\sqrt{y}-\mu)^2}{2\sigma^2}\right] \quad と \quad (2\sigma\sqrt{2\pi y})^{-1} \exp\left[-\frac{(-\sqrt{y}-\mu)^2}{2\sigma^2}\right]$$

に変換される．したがって y の全体の密度関数は (3. 2) を用いて

$$(2\sigma\sqrt{2\pi y})^{-1} e^{-(y+\mu^2)/2\sigma^2}[e^{\mu\sqrt{y}/\sigma^2} + e^{-\mu\sqrt{y}/\sigma^2}]$$
$$= (\sigma\sqrt{2\pi y})^{-1} e^{-(y+\mu^2)/2\sigma^2} \cosh\frac{\mu\sqrt{y}}{\sigma^2} \qquad (3a. 1. 16)$$

となる．$\mu = 0$ ならば y の密度関数は

$$G\left(y \,\middle|\, \frac{1}{2\sigma^2}, \frac{1}{2}\right) = \frac{1}{\sigma\sqrt{2\pi}} e^{-y/2\sigma^2} y^{-1/2}, \qquad y > 0$$
$$= 0, \qquad y \leqslant 0 \qquad (3a. 1. 17)$$

となる．これは理論物理学でレイリー分布として知られている．

これはその性質を 3a. 2 で論じる，ガンマ分布 $G(\alpha, p)$ と呼ばれる一般的な密度関数 $G(y|\alpha, p) = c \cdot e^{-\alpha y} y^{p-1}$ の特別な場合である．

3a. 2　ガンマ分布

一般的なガンマ分布は密度関数

$$G(x|\alpha, p) = \frac{\alpha^p}{\Gamma(p)} e^{-\alpha x} x^{p-1}, \quad \alpha > 0, \ p > 0, \ 0 < x < \infty \qquad (3a. 2. 1)$$

をもつ．

(i)　原点まわりの r 次の積率は $\Gamma(p+r)/\alpha^r\Gamma(p)$ であることがわかるから

$$E(x) = p/\alpha, \qquad V(x) = p/\alpha^2.$$

(ii)　$x_i \sim G(\alpha, p_i)$,　$i = 1, \cdots, k$ はすべて独立とする．このとき

$$x_1 + \cdots + x_k \sim G(\alpha, p = \sum p_i) \qquad (3a. 2. 2)$$

である．つまりガンマ分布は正規分布のように再生性をもつ．しかし2つの母数に変動があるときは成り立たない．

x_j の特性関数は

$$\int e^{itx_j} G(x_j|\alpha, p_j) dx_j = \left(\frac{\alpha}{\alpha - it}\right)^{p_j} \qquad (3a. 2. 3)$$

である．このとき $x_1 + \cdots + x_k$ の特性関数は $[\alpha/(\alpha - it)]^p$ であり，上の結果が成立する．すべての x_i について母数 α が同じでないならばこの結果は正しくない．

(iii)　$x \sim G(\alpha, p_1)$ と $y \sim G(\alpha, p_2)$ は独立とする．このとき $r = x + y$ と $f = x/y$ は独立な分布に従う．r の分布は $G(\alpha, p_1 + p_2)$ であり，f の分布は (3a. 2. 6) で定義するように $f(2p_1, 2p_2)$ である．

x と y の同時分布は

$$c \cdot e^{-\alpha x - \alpha y} x^{p_1 - 1} y^{p_2 - 1} dx dy \qquad (3a. 2. 4)$$

である．$x = r\sin^2\theta$, $y = r\cos^2\theta$, $(0 < r < \infty,\ 0 < \theta < \pi/2)$ なる変換を行うと $dxdy = 2r\sin\theta\cos\theta dr d\theta$ となる．(3a. 2. 4) は

$$c \cdot e^{-\alpha r} r^{p_1 + p_2 - 1} (\sin\theta)^{2p_1 - 1} (\cos\theta)^{2p_2 - 1} dr d\theta$$

に変換され，これは r と θ すなわち r と f が独立な分布に従うことを示している．r の分

布は $G(\alpha, p_1+p_2)$ であり，θ のみの分布は定数を付け加えると

$$\frac{2\Gamma(p_1+p_2)}{\Gamma(p_1)\Gamma(p_2)}(\sin\theta)^{2p_1-1}(\cos\theta)^{2p_2-1}d\theta, \quad 0<\theta<\frac{\pi}{2} \tag{3a.2.5}$$

となる．さて

$$f=\frac{x}{y}=\frac{\sin^2\theta}{\cos^2\theta}=\tan^2\theta,$$

$$df=2\tan\theta\sec^2\theta d\theta=2(\sin\theta/\cos^3\theta)d\theta$$

である．これを (3a.2.5) に代入すると2つの独立なガンマ確率変数の比である f の密度関数は

$$\frac{\Gamma(p_1+p_2)}{\Gamma(p_1)\Gamma(p_2)}\frac{f^{p_1-1}}{(1+f)^{p_1+p_2}} \tag{3a.2.6}$$

となる．ガンマ分布を特徴づける命題（ⅲ）の逆は Khatri and Rao (1968b) によって与えられている．

（ⅳ）$x\sim G(\alpha, p_1)$ と $y\sim G(\alpha, p_2)$ は独立とする．このとき $g=x/(x+y)$ の密度関数は

$$\frac{\Gamma(p_1+p_2)}{\Gamma(p_1)\Gamma(p_2)}g^{p_1-1}(1-g)^{p_2-1}, \quad 0<g<1 \tag{3a.2.7a}$$

であり，これはベータ分布 $B(p_1, p_2)$ と呼ばれる．

(3a.2.7a) の結果は (3a.2.5) で $g=\sin^2\theta$ を代入すれば得られる．

$p_1=p_2=\frac{1}{2}$ である特別な場合，ベータ分布は，密度関数が

$$\pi^{-1}g^{-1/2}(1-g)^{-1/2} \tag{3a.2.7b}$$

で，分布関数が

$$2\pi^{-1}\sin^{-1}g^{1/2} \tag{3a.2.7c}$$

となる．このため (3a.2.7b) は逆正弦密度関数と呼ばれる．

（ⅴ）**待ち時間の分布**　2つの互いに重なり合わない時間間隔内では事象の生起が独立であるという条件に従うとき，ある時間での事象の生起を考えよう．時間間隔 t 内で，ある事象が一回も起こらない確率を $q(t)$ とする．t_1 と t_2 が2つの連続した時間間隔ならば，独立性の仮定により $q(t_1)\cdot q(t_2)=q(t_1+t_2)$ を得る．これは q が連続ならば $q(t)=e^{-\lambda t}$ という形の解をもつ．したがって事象が起こるまでの待ち時間の分布関数は $1-e^{-\lambda t}$ である．密度関数は

$$\frac{d}{dt}(1-e^{-\lambda t})=\lambda e^{-\lambda t}=G(t|\lambda, 1) \tag{3a.2.8}$$

となり，これはガンマ分布 (3a.2.1) の特別な場合である．事象が k 回起こるまでの待ち時間の分布が $G(\lambda, k)$ であることは容易に導くことができる．

ある事象が起こるまでの待ち時間の分布が (3a.2.8) であるとしよう．一定の時間 t 内で事象が n 回起こる確率 $P(n, t)$ はいくらか？　$t_1, \cdots, t_n$ を時間間隔とし，そこで事象が n 回起こるとする．さらに $\sum t_i$ から t までの時間では事象は起こらないとする．このとき t までに事象が n 回起こる確率は領域 $R: t_1+\cdots+t_n\leqslant t$ の上で

$$\int_R \lambda^n e^{-\lambda(t_1+\cdots+t_n)}e^{-\lambda(t-t_1-\cdots-t_n)}dt_1\cdots dt_n=\int_R \lambda^n e^{-\lambda t}dt_1\cdots dt_n$$

$$=\lambda^n e^{-\lambda t}\int_R dt_1\cdots dt_n$$

となる．R の上で積分すると $P(n, t)=e^{-\lambda t}(\lambda t)^n/n!$ であることがわかる．これはポアソン分布の確率である．

カイ2乗分布　ガンマ分布の特別な場合で，$\alpha=\frac{1}{2}$ で $p=k/2$，ただし k は整数，のとき，自由度 k のカイ2乗分布と呼ばれる．密度関数は

$$\chi^2(x|k)=\frac{1}{2^{k/2}\Gamma(k/2)}e^{-x/2}x^{(k/2)-1} \tag{3a.2.9}$$

であり，これより以下の結果が得られる．

（ⅵ）$x_i\sim\chi^2(k_i)$，$i=1, \cdots, m$ は独立とすると

$$\sum x_i\sim\chi^2(\sum k_i) \tag{3a.2.10}$$

である．

（ⅶ）$x_1\sim\chi^2(k_1)$，$x_2\sim\chi^2(k_2)$ は互いに独立とする．このとき $f=x_1/x_2$ の密度関数は

$$f(x|k_1, k_2)=\frac{\Gamma\left(\frac{k_1+k_2}{2}\right)}{\Gamma\left(\frac{k_1}{2}\right)\Gamma\left(\frac{k_2}{2}\right)}\frac{x^{(k_1/2)-1}}{(1+x)^{(k_1+k_2)/2}}, \tag{3a.2.11}$$

$F=(x_1/k_1)/(x_2/k_2)$ の密度関数は

$$F(x|k_1, k_2)=\frac{\left(\frac{k_1}{k_2}\right)^{k_1/2}}{\beta\left(\frac{k_1}{2}, \frac{k_2}{2}\right)}\frac{x^{(k_1/2)-1}}{\left(1+\frac{k_1x}{k_2}\right)^{(k_1+k_2)/2}}, \tag{3a.2.12}$$

$z=(1/2)\log F$ の密度関数は

$$z(x|k_1, k_2)=\frac{2\left(\frac{k_1}{k_2}\right)^{k_1/2}}{\beta\left(\frac{k_1}{2}, \frac{k_2}{2}\right)}\frac{e^{k_1x}}{\left(1+\frac{k_1}{k_2}e^{2x}\right)^{(k_1+k_2)/2}}, \tag{3a.2.13}$$

$g=x_1/(x_1+x_2)$ の密度関数は

$$B\left(g\,\middle|\,\frac{k_1}{2}, \frac{k_2}{2}\right)=\frac{\Gamma\left(\frac{k_1+k_2}{2}\right)}{\Gamma\left(\frac{k_1}{2}\right)\Gamma\left(\frac{k_2}{2}\right)}g^{(k_1/2)-1}(1-g)^{(k_2/2)-1}. \tag{3a.2.14}$$

記号のちがいを除けば (3a. 2. 10) は (3a. 2. 2) と同じであり，(3a. 2. 11) は (3a. 2. 6) と同じである．(3a. 2. 12) の結果は (3a. 2. 11) で変数 f を $F = fk_2/k_1$ に変換すれば得られる．(3a. 2. 12) で変数 F を $z = 2^{-1}\log F$ に変換すれば (3a. 2. 13) が得られる．(3a. 2. 14) は (3a. 2. 7a) と同じである．

(3a. 2. 12) の分布はフィッシャーの分散比の分布と呼ばれ，(3a. 2. 13)の分布はフィッシャーの z 分布 (Fisher (1924)) である．これらの分布は有意性検定において重要である．

(viii) $\chi^2(k)$ の特性関数は $(1-2it)^{-k/2}$ である．これは (3a. 2. 3) で $\alpha = 1/2$, $p = k/2$ を代入すれば得られる．

$\chi^2(k)$ の4次までの積率は次のとうりである．

$$\mu_1' = k \text{ (自由度)}, \quad \mu_2 = 2k, \quad \mu_3 = 8k, \quad \mu_4 = 48k + 12k^2$$

3a. 3　ベータ分布

ベータ分布は2つの母数 γ, δ を含み，密度関数は

$$B(x|\gamma,\delta) = [\beta(\gamma,\delta)]^{-1}x^{\gamma-1}(1-x)^{\delta-1}, \qquad 0 < x < 1 \tag{3a. 3. 1}$$

である．ここで $\beta(\gamma,\delta)$ はベータ関数で $\Gamma(\gamma)\Gamma(\delta)/\Gamma(\gamma+\delta)$ である．

(i) 原点まわりの s 次の積率は

$$[\beta(\gamma,\delta)]^{-1}\ x^{s+\gamma-1}(1-x)^{\delta-1}dx = \beta(s+\gamma,\delta)/\beta(\gamma,\delta) \tag{3a. 3. 2}$$

である．したがって $E(x) = \gamma/(\gamma+\delta)$, $V(x) = \gamma\delta/(\gamma+\delta)^2(\gamma+\delta+1)$ である．

(ii) x, y がそれぞれ母数 (γ_1,δ_1), (γ_2,δ_2) の独立なベータ確率変数ならば $\gamma_1 = \gamma_2 + \delta_2$ のとき $u = xy$ の分布もまた母数 $(\gamma_2, \delta_1+\delta_2)$ のベータ分布である．

これを証明するのに，変換 $u = xy$, $v = x$ を行う．(x, y) 平面の領域 $(0 < x < 1, 0 < y < 1)$ は (u, v) 平面の領域 $(u < v < 1,\ 0 < u < 1)$ に変換される．ヤコビアンは $D(x,y)/D(u,v) = -1/v$ であるから

$$\begin{aligned} &c\cdot x^{\gamma_1-1}(1-x)^{\delta_1-1}y^{\gamma_2-1}(1-y)^{\delta_2-1}dxdy \\ &\qquad = c\cdot(1-v)^{\delta_1-1}u^{\gamma_2-1}(v-u)^{\delta_2-1}dudv \end{aligned}$$

となる．v について u から1まで積分すると u の分布は

$$c\cdot u^{\gamma_2-1}(1-u)^{\delta_1+\delta_2-1}du \tag{3a. 3. 3}$$

となる．

(iii) 一般に $\gamma_i = \gamma_{i+1} + \delta_{i+1}$, $i = 1, \cdots, k-1$ となるような母数 $(\gamma_1,\delta_1), \cdots, (\gamma_k,\delta_k)$ の k 個のベータ確率変数の積も母数 $(\gamma_k, \delta_1 + \cdots + \delta_k)$ のベータ確率変数である．

(iv) ベータ分布の特別な場合に長方形分布 $B(x|1,1)dx = dx$ がある．この密度関数は1である．確率微分 $f(u)du$ をもつ確率変数 u を考え，u の分布関数をあらわす

$$x = \int_{-\infty}^{u} f(u)du \tag{3a. 3. 4}$$

で定義された確率変数 x の分布を求めれば，長方形分布は自然に得られる．$dx = f(u)du$ だから x の密度関数は簡単に1であることがわかる．

(v) **ピアソンの p_λ 統計量**　x は長方形分布に従うとする．$y = -2\log_e x$ とおくと，$dy = (-2/x)dx$ である．したがって，y の分布は $2^{-1}e^{-y/2}dy$ となり，これは $G(\frac{1}{2}, 1)$ であり，(3a. 2. 9)で定義した $\chi^2(2)$ である．$x_1, \cdots, x_k$ を異ってもよい k 個の母集団からの k 個の独立な観測値 $u_1, \cdots, u_k$ に関係した k 個の確率 (3a. 3. 4) とすると，それぞれ自由度2の k 個の独立なカイ2乗確率変数の和である統計量

$$p_\lambda = -2\log_e x_1 - \cdots - 2\log_e x_k \tag{3a. 3. 5}$$

はそれ自身自由度 $2k$ のカイ2乗分布に従う．

この分布は，統計的仮説のいくつかの独立な有意性検定の結果を結合するのに用いられる．

3a. 4　コーシー分布

α 粒子が放射している放射線源を考え，2次元のスクリーンに衝突する粒子の分布を求めよう．スクリーンは放射線源から垂直距離 δ のところにあるとする．スクリーンに平行な平面上に放射線源を通る直交軸 (x', y') を考える．α 粒子が放射する方向は2つの座標——その方向が (x', y') 平面の垂線となす角 ϕ とその方向の (x', y') 平面への射影が x' 軸となす角 θ ——によって規定される．ここで“α 粒子の放射方向はランダムである”ということと，“放射線源をおおう球面上で面積が等しければ α 粒子を受けとる確率は等しい”ということは同等であると仮定する．つまり，いいかえると粒子が球面上の $(\theta, \theta+d\theta)$, $(\phi, \phi+d\phi)$ で定められた領域に衝突する確率は

$$c\cdot\sin\phi\, d\theta d\phi \tag{3a. 4. 1}$$

であり，これは球面の面積要素である．われわれが求めたいのは任意の原点をもつ(直交)軸に関してスクリーン上に粒子が衝突する点の座標 (x, y) の同時分布である．(x, y) 軸に関して，放射線源からスクリーンへの垂線の足の座標を (μ_1, μ_2) としよう．このとき x, y 軸は x', y' 軸に平行であると仮定すると次の関係 (δ を垂線の長さとする) を得る．

$$\begin{aligned} x = \mu_1 + \delta\tan\phi\cos\theta, \qquad y = \mu_2 + \delta\tan\phi\sin\theta, \\ dxdy = \delta^2\sin\phi\cos^{-3}\phi d\theta d\phi. \end{aligned} \tag{3a. 4. 2}$$

これを (3a. 4. 1) に代入すると2変量コーシー分布と呼ばれる (x, y) の同時分布

$$c\cdot[\delta^2 + (x-\mu_1)^2 + (y-\mu_2)^2]^{-3/2}dxdy \tag{3a. 4. 3}$$

が得られる．これを y について積分すると範囲 $(-\infty, \infty)$ の x のみの周辺密度関数は，

μ_1 の代わりに μ と書くと

$$C(x|\mu,\delta)=\frac{c\,dx}{\delta^2+(x-\mu)^2},\qquad c=\frac{\delta}{\pi} \tag{3a.4.4}$$

となる．これは位置母数 μ，尺度母数 δ をもつ1次元コーシー分布として知られている．1次元の障壁と，そこから垂直距離 δ だけ離れた点源を考え，α 粒子が障壁に衝突する点の，任意の原点から測った距離を変数にとったとき (3a.4.4) と同じ分布が得られる．

(3a.4.4) の分布はモードが $x=\mu$ である対称な分布である．x の期待値は

$$E(x)=\frac{\delta}{\pi}\int_{-\infty}^{\infty}\frac{x\,dx}{\delta^2+(x-\mu)^2}$$

である．しかしこれは主値

$$\lim_{m\to\infty}\frac{\delta}{\pi}\int_{-m}^{+m}\frac{x\,dx}{\delta^2+(x-\mu)^2}=\mu$$

をもつが，積分は存在しない．2次のモーメントは無限である．このように平均と分散が存在しない連続分布の例が得られた．

(3a.4.4) の特性関数は

$$e^{it\mu}e^{-|t\delta|} \tag{3a.4.5}$$

である．位置母数が $\mu_1,\cdots,\mu_n$ で，尺度母数は同じ δ である n 個のコーシー分布の母集団から得られる n 個の独立な観測値の平均値の特性関数は

$$e^{it\mu}e^{-|t\delta|} \tag{3a.4.6}$$

である．ただし $\mu=(\mu_1+\cdots+\mu_n)/n$．したがって n 個の観測値の平均値もまたコーシー分布に従い，その位置母数は個々の位置母数の平均値である．同じコーシー分布の母集団から得られる観測値の任意の個数の平均値は，ただ1個の観測値と同じ分布に従うことがわかる．

このように，観測値の数が増加したときの平均値の性質を述べている 2c.3 の大数の法則はコーシー分布の場合にはあてはまらない．

3a.5　スチューデントの t 分布

(i)　$y\sim N(0,1)$ と $x\sim\chi^2(k)$ は独立な確率変数とする．Student (1908) は正規確率変数とそれと独立なカイ2乗確率変数をその自由度で割ったものの平方根との比である統計量

$$t=\frac{y}{(x/k)^{1/2}} \tag{3a.5.1}$$

を定義した．y と x の同時分布は

$$c\cdot e^{-y^2/2}e^{-x/2}x^{(k/2)-1}\,dy\,dx \tag{3a.5.2}$$

である．

$$y=r\sin\theta,\quad x=r^2\cos^2\theta,\quad dx\,dy=2r^2\cos\theta\,dr\,d\theta \tag{3a.5.3}$$

により極座標 $(0<r<\infty,\ -\pi/2<\theta<\pi/2)$ に変換すると，(3a.5.2) は

$$c\cdot e^{-r^2/2}r^k(\cos\theta)^{k-1}dr\,d\theta \tag{3a.5.4}$$

に変換される．θ のみの分布は r にはよらず，

$$c\cdot(\cos\theta)^{k-1}d\theta=\left[\beta\left(\frac{1}{2},\frac{k}{2}\right)\right]^{-1}(\cos\theta)^{k-1}d\theta \tag{3a.5.5}$$

となる．ここで定数は全積分が1になるようにきめられる．求めるべき分布に従う統計量は

$$t=\frac{\sqrt{k}\,y}{\sqrt{x}}=\sqrt{k}\tan\theta\qquad(-\infty<t<\infty)$$

であり

$$dt=\sqrt{k}\sec^2\theta\,d\theta=\sqrt{k}\left(1+\frac{t^2}{k}\right)d\theta$$

である．式 (3a.5.5) は

$$S(t|k)dt=\left[\sqrt{k}\,\beta\left(\frac{1}{2},\frac{k}{2}\right)\right]^{-1}\left(1+\frac{t^2}{k}\right)^{-(k+1)/2}dt \tag{3a.5.6}$$

に変換される．これは自由度 k のスチューデントの t 分布と呼ばれ，$S(k)$ であらわす．

(ii)　$y\sim N(\mu,\sigma^2)$ と $(x/\sigma^2)\sim\chi^2(k)$ は独立とする．このとき

$$\frac{y-\mu}{(x/k)^{1/2}}\sim S(k) \tag{3a.5.7}$$

である．

(3a.5.7) の結果は，$\dfrac{y-\mu}{\sigma}\sim N(0,1)$ に注意し，(3a.5.6) を $(y-\mu)/\sigma$ と $(x/k\sigma^2)^{1/2}$ の比に適用すれば得られる．

(iii)　$y\sim N(\mu,\sigma^2)$ と $(x/\sigma^2)\sim\chi^2(k)$ は独立とする．このとき $t=y/(x/k)^{1/2}$ の密度関数は $\delta=\mu/\sigma$ とすると

$$S(t|k,\delta)=\frac{1}{\sqrt{\pi}}\frac{k^{k/2}}{\Gamma(k/2)}\frac{e^{-\delta^2/2}}{(k+t^2)^{(k+1)/2}}\sum_{s=0}^{\infty}\Gamma\left(\frac{k+s+1}{2}\right)\frac{(t\delta)^s}{s!}\left(\frac{2}{k+t^2}\right)^{s/2} \tag{3a.5.8}$$

である．これは非心 t 分布と呼ばれる．

y と x の同時分布は

$$c_1e^{-(y-\mu)^2/2\sigma^2}e^{-x/2\sigma^2}x^{(k/2)-1}dy\,dx$$

である．ただし $c_1^{-1}=\sqrt{2\pi}\sigma^k2^{k/2}\Gamma(k/2)$ である．(3a.5.3) と同様な変換 $y=r\sin\theta$, $x=r^2\cos^2\theta$ を行えば，r と θ の同時密度関数は

$$c_2e^{-(r^2-2\mu r\sin\theta)/2\sigma^2}r^k(\cos\theta)^{k-1}dr\,d\theta \tag{3a.5.9}$$

であることがわかる．ただし $c_2 = 2c_1 e^{-\mu^2/2\sigma^2} = 2c_1 e^{-\delta^2/2}$ である．これは r と θ が (3a. 5. 4)のようには独立でないことを示している．$\exp(\mu r \sin\theta/\sigma^2)$ を展開し (3a. 5. 9) を無限級数で書くと

$$c_2 e^{-r^2/2\sigma^2}(\cos\theta)^{k-1} \sum \frac{(\mu \sin\theta)^s r^{k+s}}{s!\sigma^{2s}} \tag{3a. 5. 10}$$

となる．(3a. 5. 10) で r に関して項別積分すると θ の密度関数

$$\sigma^{k+1}\ {}^{-1}c_2(\cos\theta)^{k-1} \sum \Gamma\left(\frac{k+s+1}{2}\right) 2^{(s+k+1)/2} \frac{(\delta \sin\theta)^s}{s!} \tag{3a. 5. 11}$$

が得られる．一方 $t = \sqrt{k}\tan\theta$ である．θ から t に変換し (3a. 5. 9) の c_2 の値を用いれば結局 (3a. 5. 8) のような t の分布が得られる．

非心ベータ分布と非心 f 分布はこの章の終りの問題 17 で導く．

3a. 6　統計力学での平衡状態をあらわす分布

座標系 $x_1, \cdots, x_m$ をもつ一般的な空間（たとえば物理学での相空間）を考え，点$(x_1, \cdots, x_m)$ のまわりの微小体積 Δv 中の（分子のような）粒子の数と Δv の比の極限で密度関数 $p(x_1, \cdots, x_m)$ を定義する．粒子の分布が，運動の制約に矛盾しないで空間全体にできるかぎり均等に分布しているならばその系は平衡であるという．たとえば位置 $(x_1, \cdots, x_m)$ を占める粒子はエネルギー $E(x_1, \cdots, x_m)$ をもつであろう．1つの課せられた制約は粒子1個当たりの平均エネルギーが，与えられた定数 c に等しいことである．これは

$$\int E p dv = c \tag{3a. 6. 1}$$

という関係であらわされる．ただし dv は体積要素をあらわす．われわれの目的は条件 (3a. 6. 1) の下でできるかぎり一様分布に近い p をきめることである．このためには与えられた分布の一様分布への近さをあらわす指標が必要である．一様分布に近くなればより大きな値をもつ，1つのこのような尺度

$$-\int p \log p\, dv \tag{3a. 6. 2}$$

は系のエントロピーと呼ばれる．このとき問題は (3a. 6. 1) の条件の下で (3a. 6. 2) を最大にすることである．

（i）$\int E p dv = c$ の下で $-\int p \log p dv$ が最大となるような最適な p を選ぶと，

$$p = e^{\lambda E + \mu} = \alpha e^{\lambda E} \tag{3a. 6. 3}$$

となる．ここで α は $\int p dv = 1$ となるように選ばれる．

(1e. 6. 6) で任意の2つの密度関数についての不等式

$$\int p \log \frac{p}{q} dv \geqslant 0 \tag{3a. 6. 4}$$

を証明した．これは，$\log q = \lambda E + \mu$ を選び，(3a. 6. 1) と p の積分が1であることを使えば

$$-\int p \log p dv \leqslant -\int p \log q dv = -\int p(\lambda E + \mu) dv = -(\lambda c + \mu) \tag{3a. 6. 5}$$

であることを意味している．

$$\int q dv = 1 = \int e^{\lambda E + \mu} dv, \qquad \int E q dv = c = \int E e^{\lambda E + \mu} dv$$

となるような λ と μ が存在するとしよう．このとき (3a. 6. 5) から $-\lambda c - \mu$ は(3a. 6. 2)の一定の上限となる．この上限は $p = \exp(\lambda E + \mu) = \alpha \exp(\lambda E)$ を選べば得られる．

解 (3a. 6. 3)はマックスウェル＝ボルツマン分布として知られている．定数 λ は $-1/kT$ と書ける．ここで T は温度．k はボルツマン定数である．次に2，3の応用を考えよう．

沈降平衡にたいするヘリーの法則　質量 m の粒子のエネルギーは位置座標にのみ依存すると仮定する．たとえば，粒子が地球上の重力場にあるとし，高さ z でポテンシャルエネルギー $E = mgz$ をもつとする．(3a. 6. 3) の分布によれば平衡状態で高さ z と $z + dz$ の間にある粒子の数は

$$A e^{-mgz/kT} dz \tag{3a. 6. 6}$$

である．このように粒子が液体中に浮遊していれば，粒子の数は高さとともに指数的に減少するであろう．このことは実際に観測されている．

磁場における微小な磁石の軸の角度の分布　磁場 H でモーメント p をもつ微小な磁石のポテンシャルエネルギーは

$$E = -pH\cos\theta$$

で与えられる．ただし θ は双極子の軸が磁場となす角である．この場合の空間は双極子の軸の方向余弦をあらわす球の表面である．球面の任意の点での密度は pH/kT を κ と書くと

$$A e^{pH\cos\theta/kT} = A e^{\kappa\cos\theta} \tag{3a. 6. 7}$$

となる．極座標では表面の要素は $\sin\theta d\theta d\phi$ となることに注意すると定数 A は

$$1 = \int_0^\pi \int_0^{2\pi} A e^{\kappa\cos\theta} \sin\theta d\theta d\phi = A \frac{4\pi \sinh\kappa}{\kappa} \tag{3a. 6. 8}$$

という関係から求められる．したがって (3a. 6. 7) の密度は

$$\frac{\kappa}{4\pi \sinh\kappa} e^{\kappa\cos\theta} \tag{3a. 6. 9}$$

と書ける．この分布は Fisher (1953), Watson (1956), Watson and Williams (1956) らによって岩石中の残留磁気の方向について標本を分析する際に用いられた．

マックスウェルの速度分布　理想気体における〔質量 m の〕粒子の直交する3方向の位置座標を x, y, z とし，速度を u, v, w とする．座標が (x, y, z, u, v, w) である粒子の運動エネルギーは $E = m(u^2 + v^2 + w^2)/2$ である．これより粒子の最適な分布 (3a. 6. 3) に

相当する確率微分は

$$Ae^{-\frac{m(u^2+v^2+w^2)}{2kT}}dudvdwdxdydz$$

である．x, y, z について積分すると速度のみの分布は

$$Ae^{-\frac{m(u^2+v^2+w^2)}{2kT}}dudvdw \tag{3a.6.10}$$

である．これは3方向の速度が独立な分布に従い，各成分は正規分布に従うことを示している．(3a.6.3) のより進んだ応用については，読者は Joos の『理論物理学』(1934) を参照せよ．

(3a.6.2) とは別の

$$\frac{1}{1-\alpha}\log\int p^\alpha dv, \qquad \alpha > 0 \tag{3a.6.11}$$

のような一様分布への近さをあらわす尺度を考えることができる．条件 (3a.6.1) の下で (3a.6.11) を最大にする p を求めよう．(3a.6.11) のかわりに，

$$p^\alpha dv \tag{3a.6.12}$$

を選び，$\alpha < 1$ あるいは $\alpha > 1$ に応じて，最大あるいは最小にすることができる．被積分関数に平均値の定理を用いると

$$\int(p^\alpha - q^\alpha)dv = \int(p-q)\alpha q^{\alpha-1}dv + \frac{1}{2}\int\alpha(\alpha-1)(p-q)^2 r^{\alpha-2}dv \tag{3a.6.13}$$

と書ける．ただし $r \in (p, q)$ である．条件 (3a.6.1) から λ と μ が

$$\int q dv = 1 = \int(\lambda E + \mu)^{1/(\alpha-1)}dv, \qquad \int Eq dv = c = \int E(\lambda E + \mu)^{1/(\alpha-1)}dv$$

をみたすような $q^{\alpha-1} = (\lambda E + \mu)$ を選ぼう．これを $q^{\alpha-1}$ に代入するとこの q の特定な関数にたいし (3a.6.13) の右辺は $\alpha \gtrless 1$ に応じて

$$\frac{1}{2}\int\alpha(\alpha-1)(p-q)^2 r^{\alpha-2}dv \gtrless 0$$

となる．したがって $\alpha \lessgtr 1$ に応じて $\int p^\alpha dv \lessgtr \int q^\alpha dv$ となる．

このようにして，$\int p^\alpha dv$ の極値は

$$p = (\lambda E + \mu)^{1/(\alpha-1)} \tag{3a.6.14}$$

のとき得られる．これは (3a.6.3) とは別の平衡分布の族を与える．実際，物理的な場でのこれらの分布の適切さは調べる価値がある．

3a.7　円周上の分布

観測される基本の変数が2次元における方向，つまり簡単にいえば，角度であるような場合が実際には多くある．このような場合，標本空間は単位円の円周の点からなり，円周上のボレル集合に関して確率測度を考えなければならない．このような分布は円周分布と呼ばれる．分布が単位円の円周上のルベーグ測度に関して絶対連続のとき，その分布は

$$f(x) \geqslant 0, \qquad \int_0^{2\pi} f(x)dx = 1, \qquad 0 < x \leqslant 2\pi \tag{3a.7.1}$$

をみたす周期関数（周期 2π）である密度関数 $f(x)$ で規定される．ここで x はあるきめた方向から測った角度である．

円周分布の特別な場合を論じる前に，一般的にこの分布のいくつかの特殊性を考えよう．x をあるきめた方向から測った角度で，y をその方向から角度 c にある別の方向から測った角度とする．すると

$$\begin{aligned} x &= c + y, \qquad 0 < y \leqslant 2\pi - c \\ x &= y - 2\pi + c, \qquad 2\pi - c < y \leqslant 2\pi \end{aligned} \tag{3a.7.2}$$

となり，$E(x) = c + E(y)$ は無限直線上の平行移動の場合のように成り立つ必要はない．このことは $E(x)$ が何ら“好みの方向”つまり円周上の分布の中心値をあらわしていないことを示している．

このように，もし鳥の飛ぶ方向についての独立な観測値が（鳥の移住の研究で）得られたとしても，その角度の平均値は鳥の生息地の方向の意味のある推定値になりえない．

同様に (3a.7.1) の x の分散は原点の移動に関して不変でないことが示される．このように位置と広がりに関して通常用いる尺度は円周分布からのデータを要約する際には意味をもたない．

しかしこのような場合，より意味のある尺度はつくることができる．円周上の点を角度 x に対応する1対の方向余弦 $(\cos x, \sin x)$ であらわそう．x と y が (3a.7.2) で定義されたものならば $\cos x = \cos(c + y)$，$\sin x = \sin(c + y)$ であり

$$E(\cos x) = E(\cos \overline{c + y}), \qquad E(\sin x) = E(\sin \overline{c + y}) \tag{3a.7.3}$$

となる．これより容易に

$$\cot^{-1}\frac{E(\cos x)}{E(\sin x)} = c + \cot^{-1}\frac{E(\cos y)}{E(\sin y)} \tag{3a.7.4}$$

が得られる．したがって好みの方向（角度）の意味のある尺度は

$$\cot^{-1}\frac{E(\cos x)}{E(\sin x)} \tag{3a.7.5}$$

である．(3a.7.3) からまた，

$$R^2 = [E(\cos x)]^2 + [E(\sin x)]^2 = [E(\cos y)]^2 + [E(\sin y)]^2 \tag{3a.7.6}$$

が c にはよらないことがわかる．さらに $0 \leqslant R^2 \leqslant 1$ であり，その値は x が円周上の一様分布に従うとき（広がりが最大）0であり，x が1点に集中するとき（広がりが最小）1である．したがって広がりの尺度として $\sqrt{1 - R^2}$ を使うことができる．

このように $x_1, \cdots, x_n$ が任意の角度0の方向から観測された n 個の角度ならば，好みの方向は

$$x = \cot^{-1}\frac{\sum \cos x_i}{\sum \sin x_i} \tag{3a.7.7}$$

で推定でき，広がりは $\sqrt{1-r^2}$ で推定できる，ただし

$$r^2 = \left(\frac{\sum \cos x_i}{n}\right)^2 + \left(\frac{\sum \sin x_i}{n}\right)^2. \tag{3a.7.8}$$

$\overrightarrow{OP_1}, \cdots, \overrightarrow{OP_n}$ を単位円の円周の（角度 $x_1, \cdots, x_n$ に対応する）n 個の観測点の単位ベクトルとし，$\overrightarrow{OP}$ を n 個のベクトルにふつうの平行四辺形の法則を適用して得られる合成ベクトルとする．このとき (3a.7.7) は簡単に $\overrightarrow{OP}$ の方向であり，(3a.7.8) で定義した r はそれを n で割った長さである．このように (3a.7.7) と (3a.7.8) は方向に関するデータを要約するのに合った方法である．

一様分布　基本的な円周分布として密度関数が

$$f(x) = \frac{1}{2\pi}, \qquad 0 < x \leqslant 2\pi \tag{3a.7.9}$$

である円周上の一様分布を考えることができる．この分布では，すべての方向が等しく起こりうる．好みの方向はなく (3a.7.5) の尺度は定義されない．

円周正規分布　単峰形の円周分布の中で多く利用されているのは円周正規分布である．すなわち密度関数が

$$f(x) = [2\pi I_0(\kappa)]^{-1} \exp[\kappa \cos(x-\beta)],$$
$$0 < x \leqslant 2\pi, \quad 0 \leqslant \beta \leqslant 2\pi, \quad 0 \leqslant \kappa < \infty \tag{3a.7.10}$$

である，いわゆるフォン・ミーゼスの分布である．ここで $I_0(x)$ は引数が純虚数 ix のベッセル関数である．次数が n であるより一般的な関数 $I_n(x)$ は級数

$$I_n(x) = \sum_{r=0}^{\infty} \frac{(x/2)^{n+2r}}{r!(n+r)!} \tag{3a.7.11}$$

で定義される．(3a.7.10) の密度関数を用いると

$$E(\cos x) = \rho \cos \beta, \qquad E(\sin x) = \rho \sin \beta, \tag{3a.7.12}$$
$$\rho = I_1(\kappa)/I_0(\kappa)$$

となる．したがって

$$\cot^{-1}\frac{E(\cos x)}{E(\sin x)} = \beta \tag{3a.7.13}$$

となり，これは β が好みの方向をあらわす母数であることを示している．広がり $\sqrt{1-\rho^2}$ は他の母数 κ にのみ依存し，κ の値が小さいほど広がりが大きくなる．したがって κ を広がりの尺度と考えることができる．

n 個の独立な観測値 $x_1, \cdots, x_n$ が得られると，β と κ は

$$\cot \beta = \frac{\sum \cos x_i}{\sum \sin x_i}, \qquad \frac{I_1(\kappa)}{I_0(\kappa)} = r \tag{3a.7.14}$$

で推定される．ただし r は (3a.7.8) で定義されている．

巻きこみ分布　“単位円のまわりに直線上の分布を巻きこむこと” によって円周分布が得られる．たとえば $N(0, \sigma^2)$ を巻きこむと，密度関数が

$$f(x) = \frac{1}{2\pi} + \frac{1}{\pi}\sum_{n=1}^{\infty} \cos nx \exp(-n^2\sigma^2/2) \tag{3a.7.15}$$

である巻きこみ正規分布を得る．同様にコーシー分布の密度関数 $\delta/\pi(\delta^2 + x^2)$，$-\infty < x < \infty$ を選ぶと密度関数が

$$f(x) = \frac{1-\sigma^2}{2\pi[1+\sigma^2-2\sigma\cos x]} \tag{3a.7.16}$$

である巻きこみコーシー分布を得る．ここで $0 < x \leqslant 2\pi$，$0 \leqslant \sigma = e^{-\delta} < 1$ である．

2峰形の分布　円周上のデータで好みの方向が1つより多いときは多峰形の円周分布を考えなければならない．たとえば

$$f(x) = [2\pi I_0(\kappa)]^{-1} \exp[\kappa \cos 2(x-\beta)] \tag{3a.7.17}$$

は，モードが β と $\pi + \beta$ である軸対称な分布の密度関数となる．もう1つの興味ある2峰形の密度関数は

$$f(x) = \frac{(1-\rho^2)^{1/2}}{\pi(1-\rho \sin 2x)} \tag{3a.7.18}$$

である．(X_1, X_2) が平均0，分散1，相関係数 ρ の2変量正規密度関数をもつならば，$x = \tan^{-1}(X_1/X_2)$ が密度関数 (3a.7.18) をもつことは容易に示される．

同様に方向が3次元で観測されるとき，球上の分布を考えることができる．このような分布の例と特徴づけはすでに (3a.6.9) で考えた．

方向に関するデータのくわしい取り扱いは Mardia(1972) に含まれている．この本の付録2.2に (3a.7.14) で与えられた r にたいし κ を求める表をのせてある．また応用については J.S. Rao and Sengupta(1972) を見よ．

3b　標　本　分　布

3b.1　定　義　と　結　果

標本分布を導くのに使われる方法は変数変換である．ふつう，線形変換 $\mathbf{Y} = \mathbf{AX}$ は $\mathbf{X}$

から $\mathbf{Y}$ ($\mathbf{X}, \mathbf{Y}$ は同じ大きさの列ベクトル) へ正則な行列 $\mathbf{A}$ を用いて行われる．変換のヤコビアン $D\mathbf{Y}/D\mathbf{X}$ は $|\mathbf{A}|$ で，これは正の符号をもつとすると，微分要素は

$$dy_1 dy_2 \cdots dy_n = |\mathbf{A}| dx_1 dx_2 \cdots dx_n \tag{3b.1.1}$$

という関係で結ばれる．これを $d\mathbf{Y} = |\mathbf{A}| d\mathbf{X}$ であらわそう．$\mathbf{A}$ が直交行列のとき $|\mathbf{A}| = 1$ である．

変換行列の2つの型が一般に用いられる．1つは完全な直交行列である．**1b.4** で示したように，$\mathbf{A}$ が直交行列ならば $\mathbf{Y} = \mathbf{AX}$ は，2次形式を簡単に

$$\mathbf{X}'\mathbf{X} \to \mathbf{Y}'\mathbf{Y}$$

$$(\mathbf{X} - \boldsymbol{\mu})'(\mathbf{X} - \boldsymbol{\mu}) \to (\mathbf{Y} - \boldsymbol{\eta})'(\mathbf{Y} - \boldsymbol{\eta}), \qquad \boldsymbol{\eta} = \mathbf{A}\boldsymbol{\mu} \tag{3b.1.2}$$

のように変換する，つまり距離を保つ．

もう1つの型は $\mathbf{A}_i\mathbf{A}_j' = 0,\ i \neq j$ である分割行列 ($\mathbf{A}_i$ が $n_i \times n$ 型で $\sum n_i = n$)

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_1 \\ \vdots \\ \mathbf{A}_k \end{pmatrix}$$

である．したがって部分行列は互いに直交するがそれ自身は必ずしも直交していない．変換 $\mathbf{Y} = \mathbf{AX}$ は部分行列を用いて

$$\mathbf{Y}_1 = \mathbf{A}_1\mathbf{X}, \quad \cdots, \quad \mathbf{Y}_k = \mathbf{A}_k\mathbf{X}$$

と書ける．ここで $\mathbf{Y}_1, \cdots, \mathbf{Y}_k$ は新しい変数の互いに排反な部分集合である．このような変換では

$$\mathbf{X}'\mathbf{X} \to \mathbf{Y}_1'\mathbf{B}_1\mathbf{Y}_1 + \cdots + \mathbf{Y}_k'\mathbf{B}_k\mathbf{Y}_k$$

$$(\mathbf{X} - \boldsymbol{\mu})'(\mathbf{X} - \boldsymbol{\mu}) \to (\mathbf{Y}_1 - \boldsymbol{\eta}_1)'\mathbf{B}_1(\mathbf{Y}_1 - \boldsymbol{\eta}_1) + \cdots + (\mathbf{Y}_k - \boldsymbol{\eta}_k)'\mathbf{B}_k(\mathbf{Y}_k - \boldsymbol{\eta}_k) \tag{3b.1.3}$$

となることが [(v), **1c.1**] で示されている．ここで $\mathbf{B}_i = (\mathbf{A}_i\mathbf{A}_i')^{-1}$，$\boldsymbol{\eta}_i = \mathbf{A}_i\boldsymbol{\mu}$ である．(3b.1.3) の結果により，変換後の式は新しい変数の互いに排反な部分集合の2次形式に分割されることがわかる．$\mathbf{A}$ が完全に直交しているとき，つまり各行が他のあらゆる行に直交するとき (3b.1.2) のように完全に分割される．

次にベクトル空間のいくつかの基礎概念を利用しよう．行列 $\mathbf{B}$ の列ベクトルで生成される線形部分空間を $\mathscr{M}(\mathbf{B})$ であらわす．容易に $\mathscr{M}(\mathbf{B}) = \mathscr{M}(\mathbf{BB}')$ であることがわかる．

$\mathbf{H}$ が階数 k の行列で $\mathscr{M}(\mathbf{H}) \subset \mathscr{M}(\mathbf{B}) = \mathscr{M}(\mathbf{BB}')$ ならば

$$\mathbf{H} = \mathbf{BB}'\mathbf{C}, \qquad k = \operatorname{rank} \mathbf{B}'\mathbf{C} \tag{3b.1.4}$$

となるような行列 $\mathbf{C}$ が存在する．

また1次方程式の集合 $\mathbf{GX} = \mathbf{F}$ は $\mathbf{F}$ がいかなるものでも，$\mathbf{G}$ の階数と $\mathbf{G}$ の行の数が等しければ解をもつ．

後の議論を簡単にするために他のいくつかの結果を記す．

$$\int_T f(x_1, \cdots, x_n) dx_1 \cdots dx_n = h(T), \qquad T = g(x_1, \cdots, x_n)$$

が変数 $x_1, \cdots, x_n$ を新しい変数の1つとして T を含む別の集合に変換し，T 以外のすべての新しい変数に関して積分し，微分 dT を省略した結果であるとする．このとき $h(T)$ は T の密度関数である．

すでに **3a** で導入した密度関数は以下の記号を用いることにする．

$$N(x|\mu, \sigma^2) = (\sigma\sqrt{2\pi})^{-1} e^{-(x-\mu)^2/2\sigma^2}, \qquad -\infty < x < \infty$$

$$G(x|\alpha, p) = [\alpha^p/\Gamma(p)] e^{-\alpha x} x^{p-1}, \qquad 0 < x < \infty$$

$$B(x|p, q) = [\beta(p, q)]^{-1} x^{p-1}(1 - x)^{q-1}, \qquad 0 < x < 1$$

$$\chi^2(x|k) = [2^{k/2}\Gamma(k/2)]^{-1} e^{-x/2} x^{(k/2)-1}, \qquad 0 < x < \infty$$

$$S(x|k) = \left[\sqrt{k}\,\beta\left(\frac{1}{2}, \frac{k}{2}\right)\right]^{-1} \left(1 + \frac{x^2}{k}\right)^{-(k+1)/2}, \qquad -\infty < x < \infty$$

$$f(x|k_1, k_2) = \left[\beta\left(\frac{k_1}{2}, \frac{k_2}{2}\right)\right]^{-1} \frac{x^{(k_1/2)-1}}{(1 + x)^{(k_1+k_2)/2}}, \qquad 0 < x < \infty$$

$$F(x|k_1, k_2) = \left[\beta\left(\frac{k_1}{2}, \frac{k_2}{2}\right)\right]^{-1} \left(\frac{k_1}{k_2}\right)^{k_1/2} \frac{x^{(k_1/2)-1}}{\left(1 + \frac{k_1 x}{k_2}\right)^{(k_1+k_2)/2}}, \qquad 0 < x < \infty$$

3a での証明から，次の結果を得る．

(a) $$\int_{x=x_1+\cdots+x_n} G(x_1|\alpha, p_1) \cdots G(x_n|\alpha, p_n) dx_1 \cdots dx_n = G(x|\alpha, p_1 + \cdots + p_n) \tag{3b.1.5}$$

(b) $$\int_{g=x_1/(x_1+x_2)} G(x_1|\alpha, p) G(x_2|\alpha, q) dx_1 dx_2 = B(g|p, q) \tag{3b.1.6}$$

(c) $$\int_{x=x_1/x_2} G(x_1|\alpha, p) G(x_2|\alpha, q) dx_1 dx_2 = f(x|2p, 2q) \tag{3b.1.7}$$

(d) $$\int_{t=x_1/\sqrt{x_2/k}} N(x_1|0, 1) \chi^2(x_2|k) dx_1 dx_2 = S(t|k) \tag{3b.1.8}$$

(e) $$\int_{x=x_1+\cdots+x_n} \chi^2(x_1|k_1) \cdots \chi^2(x_n|k_n) dx_1 \cdots dx_n = \chi^2(x|k_1 + \cdots + k_n) \tag{3b.1.9}$$

(f) $$\int_{x=x_1/x_2} \chi^2(x_1|k_1) \chi^2(x_2|k_2) dx_1 dx_2 = f(x|k_1, k_2) \tag{3b.1.10}$$

(g) $$\int_{x=x_1k_2/x_2k_1} \chi^2(x_1|k_1) \chi^2(x_2|k_2) dx_1 dx_2 = F(x|k_1, k_2) \tag{3b.1.11}$$

(h) $$\int_{g=x_1/(x_1+x_2)} \chi^2(x_1|k_1) \chi^2(x_2|k_2) dx_1 dx_2 = B\left(g \middle| \frac{k_1}{2}, \frac{k_2}{2}\right) \tag{3b.1.12}$$

(i) $\int_{r=x_1/(x_2+x_1^2)^{1/2}} N(x_1|0,1)\chi^2(x_2|k)dx_1dx_2$

$$= \Gamma\left(\frac{k+1}{2}\right)\Big/\left[\sqrt{\pi}\,\Gamma\left(\frac{k}{2}\right)\right](1-r^2)^{(k/2)-1} \qquad (3b.1.13)$$

(j) $\int_{y=x^2/\sigma^2} N(x|0,\sigma^2)dx = \chi^2(y|1) = G(y|\tfrac{1}{2},\tfrac{1}{2})$ (3b.1.14)

(k) $\int_{y=x^2/\sigma^2} N(x|\mu,\sigma^2)dx = e^{-\mu^2/2\sigma^2}\sum\frac{1}{r!}\left(\frac{\mu^2}{2\sigma^2}\right)^r G(y|\tfrac{1}{2},r+\tfrac{1}{2})$ (3b.1.15)

(j), (k) の結果は (3a.1.17), (3a.1.16) と同じである.

3b.2　正規確率変数の2乗和

$x_i \sim N(\mu_i,\sigma^2)$, $i=1,\cdots,k$ を独立な確率変数とする. $x=\sum x_i^2/\sigma^2$ の分布は何か? すべての i について $\mu_i=0$ ならば $x_i^2/\sigma^2 \sim \chi^2(1)$ である. たたみこみ (3b.1.9) での, カイ2乗分布の再生性を適用すると $x=\sum x_i^2/\sigma^2 \sim \chi^2(k)$ つまり $\sum x_i^2 \sim \sigma^2\chi^2(k)$ となる.

少なくとも1つの i について $\mu_i \neq 0$ ならば $\mathbf{B}$ の第1行が $(\mu_1/\mu,\cdots,\mu_k/\mu)$ であるような直交変換 $\mathbf{Y}=\mathbf{BX}$ を行う. ただし $\mu^2=\mu_1^2+\cdots+\mu_k^2$ である. 変換は直交であり, $(\mu_1,\cdots,\mu_k)\to(\mu,0,\cdots,0)$ だから

$$\sum(x_i-\mu_i)^2 = (y_1-\mu)^2+y_2^2+\cdots+y_k^2$$

となる. したがって, $x_1,\cdots,x_k$ の同時密度関数

$$c_1\exp[-\sum(x_i-\mu_i)^2/2\sigma^2]$$

は

$$c_1\exp\{-[(y_1-\mu)^2+y_2^2+\cdots+y_k^2]/2\sigma^2\}$$

に変換される. これは $y_1,\cdots,y_k$ が独立で $E(y_1)=\mu$, $E(y_i)=0$, $i>1$ の正規分布に従うことをあらわしている. $\mathbf{X'X}\to\mathbf{Y'Y}$ より与えられた統計量は

$$x=\frac{y_1^2}{\sigma^2}+\frac{y_2^2+\cdots+y_k^2}{\sigma^2}=u+v$$

となる. したがって $v\sim\chi^2(k-1)$ すなわち $G[\tfrac{1}{2},(k-1)/2]$ となり, u は v とは独立な分布 (3b.1.15) に従う. ゆえに v と u の同時分布は

$$e^{-\mu^2/2\sigma^2}\sum\frac{1}{r!}\left(\frac{\mu^2}{2\sigma^2}\right)^r G\left(u\middle|\frac{1}{2},r+\frac{1}{2}\right)du\,G\left(v\middle|\frac{1}{2},\frac{k-1}{2}\right)dv \qquad (3b.2.1)$$

である. (3b.1.5) を積 (3b.2.1) の級数の各項に適用すると $x=u+v=\sum x_i^2/\sigma^2$ の分布として

$$e^{-\mu^2/2\sigma^2}\sum\frac{1}{r!}\left(\frac{\mu^2}{2\sigma^2}\right)^r G\left(x\middle|\frac{1}{2},r+\frac{k}{2}\right)dx \qquad (3b.2.2)$$

を得る. これは自由度 k, 非心率 $\lambda=(\mu^2/\sigma^2)$ の非心カイ2乗分布と呼ばれ, $\chi^2(k,\lambda)$ であらわされる. 非心率は確率変数にその期待値を代入したときの2次形式の値であることに注意せよ.

$G(\alpha,p)$ の特性関数は知られているから [(ii), 3a.2 参照] $\chi^2(k,\lambda)$ の特性関数を求めるにはガンマ分布の特性関数の (3b.2.2) のような係数をもつ和を求めればよい. したがって $\chi^2(k,\lambda)$ の特性関数は

$$e^{-\lambda/2}\sum\frac{1}{r!}\left(\frac{\lambda}{2}\right)^r(1-2it)^{-r-k/2} = (1-2it)^{-k/2}e^{\lambda it/(1-2it)} \qquad (3b.2.3)$$

である. (3b.2.3) からただちに非心カイ2乗分布は k と λ に関して再生性をもつことがわかる. このように $x_i\sim\chi^2(k_i,\lambda_i)$, $i=1,\cdots,m$ で, かつすべて独立ならば

$$x=x_1+\cdots+x_m\sim\chi^2(\sum k_i,\sum\lambda_i) \qquad (3b.2.4)$$

である.

3b.3　標本平均と標本分散の同時分布

$x_i\sim N(\mu,\sigma^2)$, $i=1,\cdots,n$ はすべて独立とする. $\bar{x}=(x_1+\cdots+x_n)/n$ と $s^2=\sum(x_i-\bar{x})^2/(n-1)$ の同時分布を求めよう.

$\mathbf{B}$ の第1行が $(n^{-1/2},\cdots,n^{-1/2})$ であるような直交変換 $\mathbf{Y}=\mathbf{BX}$ を行う. すると次のことが成り立つ.

(a) $y_1=\sqrt{n}\,\bar{x}$.

(b) $y_1^2+\cdots+y_n^2=x_1^2+\cdots+x_n^2=\sum(x_i-\bar{x})^2+n\bar{x}^2$

$y_2^2+\cdots+y_n^2=(n-1)s^2$.

(c) $(y_1-\sqrt{n}\,\mu)^2+y_2^2+\cdots+y_n^2=(x_1-\mu)^2+\cdots+(x_n-\mu)^2$.

(c) により, 密度関数 $c\cdot e^{-\Sigma(x_i-\mu)^2/2\sigma^2}$ は

$$c\cdot e^{-[(y_1-\mu\sqrt{n})^2+y_2^2+\cdots+y_n^2]/2\sigma^2}$$

に変換される. したがって $y_1,\cdots,y_n$ は独立な分布に従う. よって

$$\sqrt{n}\,\bar{x}=y_1\sim N(\sqrt{n}\,\mu,\sigma^2)$$

$$\frac{(n-1)s^2}{\sigma^2}=\frac{y_2^2+\cdots+y_n^2}{\sigma^2}\sim\chi^2(n-1)$$

となり, これらは互いに独立な分布に従う. 尺度を変えると $\bar{x}$ と s^2 の同時分布は

$$N\left(\bar{x}\middle|\mu,\frac{\sigma^2}{n}\right)d\bar{x}\,\chi^2\left(\frac{\overline{n-1}s^2}{\sigma^2}\middle|n-1\right)\frac{n-1}{\sigma^2}ds^2 \qquad (3b.3.1)$$

となる.

$t=\sqrt{n}(\bar{x}-\mu)/s$ を定義すると, t は $\sqrt{n}(\bar{x}-\mu)/\sigma\sim N(0,1)$ と $s^2/\sigma^2\sim(n-1)^{-1}\chi^2(n-1)$ の平方根の比であるから, スチューデントの分布 (3b.1.8) を用いると $t\sim S(n-1)$ となる.

正規確率変数の1次関数の分布　$x_i \sim N(\mu_i, \sigma^2)$, $i=1,\cdots,n$ を n 個の独立な確率変数とする．$x_1,\cdots,x_n$ の p 個 $(p<n)$ の独立な1次関数の同時分布を求めよう．この1次関数は行列の記号を用いると $\mathbf{Z}=\mathbf{BX}$ と書ける．ただし $\mathbf{Z}'=(z_1,\cdots,z_p)$, $\mathbf{B}$ は階数 p の $p\times n$ の行列である．

$\boldsymbol{\mu}'=(\mu_1,\cdots,\mu_n)$ とする．新しい確率変数 $\mathbf{U}$ の平均値が0になるように $\mathbf{X}$ を $\mathbf{U}+\boldsymbol{\mu}$ に変換する．$\mathbf{Z}$ を $\mathbf{U}$ であらわすと $\mathbf{Z}=\mathbf{B}(\mathbf{U}+\boldsymbol{\mu})$ つまり $\mathbf{Z}-\mathbf{B}\boldsymbol{\mu}=\mathbf{BU}$ となる．そこで $\mathbf{U}$ の分布をもとにまず $\mathbf{V}=\mathbf{BU}$ の分布を求めると，$\mathbf{Z}$ の分布は $\mathbf{V}=\mathbf{BU}$ を $\mathbf{Z}-\mathbf{B}\boldsymbol{\mu}$ でおきかえれば得られることがわかる．

$\mathbf{B}$ の行が正規直交であり，したがって直交行列 $\mathbf{C}$ になるように $n-p$ 個の正規直交ベクトルを行列 $\mathbf{B}$ につけ加えることができるとしよう．$\mathbf{C}$ を用いて $(u_1,\cdots,u_n)$ から $(v_1,\cdots,v_p,v_{p+1},\cdots,v_n)$ への変換を行う．密度関数は

$$c\cdot e^{-(\Sigma u_i^2)/2\sigma^2} \to c\cdot e^{-(\Sigma v_i^2)/2\sigma^2} \tag{3b.3.2}$$

となり v_i はすべて独立で正規分布に従う．正確には $\mathbf{BU}$ の要素である部分集合 $v_1,\cdots,v_p$ の同時密度関数は

$$c\cdot e^{-(v_1^2+\cdots+v_p^2)/2\sigma^2} = c\cdot e^{-\mathbf{V}'\mathbf{V}/2\sigma^2} \tag{3b.3.3}$$

である．

もし $\mathbf{B}$ の行が正規直交でないならば，$\mathbf{DB}$ の行が正規直交となるような正則な p 次正方行列 $\mathbf{D}$ が存在する．確率変数 $\mathbf{W}=\mathbf{DV}=(\mathbf{DB})\mathbf{U}$ を考えよう．$\mathbf{DB}$ の行は正規直交だから $\mathbf{W}=(\mathbf{DB})\mathbf{U}$ の分布は (3b.3.3) を用いると

$$c\cdot e^{-\mathbf{W}'\mathbf{W}/2\sigma^2}d\mathbf{W} = (2\pi\sigma^2)^{-p/2}e^{-\mathbf{W}'\mathbf{W}/2\sigma^2}d\mathbf{W}$$

となる．したがって $\mathbf{V}=\mathbf{D}^{-1}\mathbf{W}$ の分布は ($d\mathbf{W}=|\mathbf{D}|d\mathbf{V}$ を用いれば)

$$(2\pi\sigma^2)^{-p/2}e^{-\mathbf{V}'\mathbf{D}'\mathbf{D}\mathbf{V}/2\sigma^2}|\mathbf{D}|d\mathbf{V}$$

である．一方 $\mathbf{DB}$ の行は正規直交だから $\mathbf{DBB}'\mathbf{D}'=\mathbf{I} \Rightarrow (\mathbf{BB}')^{-1}=\mathbf{D}'\mathbf{D}$, $|\mathbf{BB}'|=1/|\mathbf{D}|^2$ となる．$(\mathbf{BB}')=\mathbf{A}$ と書くと $\mathbf{V}$ の分布，したがって $\mathbf{Z}$ の分布は

$$\frac{|\mathbf{A}|^{-1/2}}{(2\pi\sigma^2)^{p/2}}e^{-\mathbf{V}'\mathbf{A}^{-1}\mathbf{V}/2\sigma^2}d\mathbf{V} = \frac{|\mathbf{A}|^{-1/2}}{(2\pi\sigma^2)^{p/2}}e^{-(\mathbf{Z}-\mathbf{B}\boldsymbol{\mu})'\mathbf{A}^{-1}(\mathbf{Z}-\mathbf{B}\boldsymbol{\mu})/2\sigma^2}d\mathbf{Z} \tag{3b.3.4}$$

となる．(3b.3.4) で $\mathbf{B}\boldsymbol{\mu}$ を任意のベクトルでおきかえ，$\mathbf{A}$ を任意の正定符号行列でおきかえた密度関数は p 変量正規密度関数と呼ばれる．

関数 $\mathbf{Z}=\mathbf{BX}$ の個数 p が n より大きい，あるいは $R(\mathbf{B})<p$ としよう．このとき $\mathbf{Z}$ の分布は何か？　上で述べた密度関数 (3b.3.4) を導く方法は $\mathbf{BB}'$ がもはや正則でないのでうまくいかない．実際，$\mathbf{Z}$ の成分は線形関係があるからルベーグ測度に関して $\mathbf{Z}$ の密度関数は存在しない．しかし $\mathbf{Z}$ の分布は特性関数

$$E(e^{i\mathbf{t}'\mathbf{Z}}) = E(e^{i\mathbf{t}'\mathbf{BX}}) = E(e^{i\mathbf{T}'\mathbf{X}}), \qquad \mathbf{T}=\mathbf{B}'\mathbf{t}$$

$$= e^{i\mathbf{T}'\boldsymbol{\mu}-\sigma^2\mathbf{T}'\mathbf{T}/2} \qquad \text{((3a.1.6) を用いて)}$$

$$= e^{i\mathbf{t}'\mathbf{B}\boldsymbol{\mu}-\sigma^2\mathbf{t}'\mathbf{BB}'\mathbf{t}/2} = e^{i\mathbf{t}'\boldsymbol{\nu}-\mathbf{t}'\boldsymbol{\Sigma}\mathbf{t}/2} \tag{3b.3.5}$$

によってうまく定義できる．この特性関数は分布が2つの母数

$$E(\mathbf{Z})=\boldsymbol{\nu}, \qquad D(\mathbf{Z})=\boldsymbol{\Sigma} \tag{3b.3.6}$$

に依存していることを示している．特性関数 (3b.3.5) によってきめられる分布を密度関数の存在の有無にかかわらず p 変量正規分布と呼び $N_p(\boldsymbol{\nu},\boldsymbol{\Sigma})$ であらわす．(3b.3.4) のように $|\boldsymbol{\Sigma}|\neq 0$ のときは密度関数が存在する．(3b.3.4) で定義した分布のいくつかの性質は 3b.6 で調べる．一方 $\boldsymbol{\Sigma}$ が特異である場合を含む一般的な場合は第8章で考える．$\boldsymbol{\Sigma}$ が特異であるときの密度関数の表現は 8a.4 で与えられる．

3b.4　2次形式の分布

n 個の独立な正規確率変数 $y_i \sim N(\mu_i, 1)$, $i=1,\cdots,n$ を考え，その2次形式の分布を論じよう．$\mathbf{Y}'=(y_1,\cdots,y_n)$, $\boldsymbol{\mu}'=(\mu_1,\cdots,\mu_n)$ とする．

(i)　**フィッシャー＝コクランの定理**　$Q_1,\cdots,Q_k$ を $\mathbf{Y}'\mathbf{Y}=Q_1+\cdots+Q_k$ となるようなそれぞれ階数が $n_1,\cdots,n_k$ である k 個の2次形式とする．すると $Q_i \sim \chi^2(n_i,\lambda_i)$ で互いに独立であるための必要十分条件は $n=\sum n_i$ である．このとき $Q_i=\mathbf{Y}'\mathbf{A}_i\mathbf{Y}$ ならば $\lambda_i=\boldsymbol{\mu}'\mathbf{A}_i\boldsymbol{\mu}$ であり，かつ，$\sum \mu_j^2=\sum \lambda_i$ となる．

階数 m $(\leqq n)$ の n 個の変数の任意の2次形式 $\mathbf{Y}'\mathbf{AY}$ は $\mathbf{Y}$ から $\mathbf{X}$ への適当な線形変換によって $\pm x_1^2 \pm \cdots \pm x_m^2$ に帰着できる，つまり $\mathbf{Y}'\mathbf{AY}$ は $\mathbf{Y}$ の m 個の1次関数の2乗を ± 1 の係数で結合したもので表現できるという，(1c.1.2) で証明した結果を用いよう．$Q_i=\mathbf{Y}'\mathbf{A}_i\mathbf{Y}$ とする．$\operatorname{rank}\mathbf{A}_i=n_i$ だから

$$Q_i = \pm(b_{11}y_1+\cdots+b_{1n}y_n)^2 \pm \cdots \pm (b_{n_i1}y_1+\cdots+b_{n_in}y_n)^2$$

となる $\mathbf{Y}$ の n_i 個の1次関数が存在する．Q_1 には n_1 個の1次関数，Q_2 には n_2 個の1次関数，… を考えよう．$n=\sum n_i$ ならば $\mathbf{BY}$ と書ける n 個の1次関数が存在する．ただし $\mathbf{B}$ は $n\times n$ の行列である．次に $\sum Q_i$ は係数が ± 1 の，n 個の1次関数の2乗の結合だから

$$\sum Q_i = \mathbf{Y}'\mathbf{B}'\boldsymbol{\Delta}\mathbf{BY}$$

が成り立つ．ここで $\boldsymbol{\Delta}$ は各対角要素が $+1$ または -1 の対角行列である．一方

$$\text{すべての } \mathbf{Y} \text{ について} \quad \mathbf{Y}'\mathbf{Y}=\sum Q_i=\mathbf{Y}'\mathbf{B}'\boldsymbol{\Delta}\mathbf{BY} \Rightarrow \mathbf{I}=\mathbf{B}'\boldsymbol{\Delta}\mathbf{B} \tag{3b.4.1}$$

である．$\mathbf{B}$ の階数は n でなければならないから $\boldsymbol{\Delta}=(\mathbf{B}')^{-1}\mathbf{B}^{-1}$ となり，$\boldsymbol{\Delta}$ は正定符号である．ゆえに $\boldsymbol{\Delta}$ は対角要素がすべて $+1$ でなければならない．したがって (3b.4.1) より $\mathbf{B}$ は直交行列である．変換 $\mathbf{X}=\mathbf{BY}$ は直交変換である．したがって $\mathbf{X}$ の成分 $(x_1,$

$\cdots, x_n$) はすべて独立で正規分布に従う．また Q_1 は $\mathbf{BY}$ の最初の n_1 個の1次関数の2乗和だから Q_1 は $x_1^2+\cdots+x_{n_1}^2$ に変換される．同様に Q_2 は $x_{n_1+1}^2+\cdots+x_{n_1+n_2}^2$ に変換され，Q_3 以下も同様である．

Q_i は独立な確率変数 x_i の排反な部分集合に依存しているから互いに独立である．さらに Q_i は n_i 個の正規確率変数の2乗和であり，(3b.2.2) より $Q_i \sim \chi^2(n_i, \lambda_i)$ すなわち $\lambda_i = \mu' \mathbf{A}_i \mu$，$i=1,\cdots,k$ の非心カイ2乗分布に従う．一方 $\mathbf{Y'Y} \sim \chi^2(n, \sum \mu_j^2)$ だから (3b.2.4) を用いて $\sum \mu_j^2 = \sum \lambda_i$ となる．このようにして条件 $n = \sum n_i$ の十分性は証明できた．必要性は明らかである．

$\mu' \mathbf{A}_i \mu = 0$ のとき，$Q_i \sim \chi^2(n_i)$ すなわち中心カイ2乗分布に従う．これは $\mu_j = 0$, $j=1,\cdots,n$ のときに成り立つ．逆にすべての Q_i が中心カイ2乗分布に従うならば，すべての j にたいして $\mu_j = 0$ である．

読者は非心カイ2乗分布に慣れていないならば，フィッシャー＝コクランの定理はすべての μ_j, λ_i を0とおいて読むことができる．

後の議論で，記号を簡単にするためにすべての j にたいし $\mu_j = 0$ とする．しかし中心カイ2乗分布 $\chi^2(k)$ のかわりに非心カイ2乗分布 $\chi^2(k, \lambda)$ を書いても行列に与えた条件や結果は有効である．λ は確率変数にその期待値を代入したときの2次形式の値であることに注意しよう．(i) で証明した一般的な結果により，すべてこれらの結果は正しい．

(ii) <u>$\mathbf{Y'AY}$ がカイ2乗分布に従うための必要十分条件は，$\mathbf{A}$ がベキ等，すなわち $\mathbf{A}^2 = \mathbf{A}$ である．このときカイ2乗分布の自由度は $\operatorname{rank}\mathbf{A} = \operatorname{trace}\mathbf{A}$ である．</u>

$$\mathbf{Y'Y} = \mathbf{Y'AY} + \mathbf{Y'}(\mathbf{I}-\mathbf{A})\mathbf{Y}$$

でかつ $\mathbf{A}^2 = \mathbf{A}$ で $\operatorname{rank}\mathbf{A} + \operatorname{rank}(\mathbf{I}-\mathbf{A}) = n$ (1b.7 の (ii) を参照) だからフィッシャー＝コクランの定理が適用できるので，十分性はただちに証明できる．

必要性を証明するために，変換 $\mathbf{Y} = \mathbf{CX}$ の下で $\lambda_1, \cdots, \lambda_m$ を $\mathbf{A}$ の0でない固有値とすると

$$\begin{aligned} \mathbf{Y'AY} &\to \mathbf{X'C'ACX} = \lambda_1 x_1^2 + \cdots + \lambda_m x_m^2 \\ \mathbf{Y'Y} &\to \mathbf{X'X} = x_1^2 + \cdots + x_n^2 \end{aligned} \tag{3b.4.2}$$

となるような直交行列 $\mathbf{C}$ が存在するのを確かめよう ((1c.3.5) を参照)．

$x_j^2 \sim \chi^2(1) = G(\frac{1}{2}, \frac{1}{2})$ だから x_j^2 の特性関数は $(1-2it)^{-1/2}$ である．したがって $\lambda_1 x_1^2 + \cdots + \lambda_m x_m^2$ の特性関数は

$$[(1-2i\lambda_1 t)(1-2i\lambda_2 t)\cdots(1-2i\lambda_m t)]^{-1/2} \tag{3b.4.3}$$

である．一方 $\mathbf{Y'AY}$ はカイ2乗分布（自由度を p とする）に従うことが仮定されており，この場合その特性関数は

$$(1-2it)^{-p/2} \tag{3b.4.4}$$

である．(3b.4.3) と (3b.4.4) を比較すると $p=m$ ですべての i にたいし $\lambda_i = 1$ となる．(3b.4.2) から $\mathbf{C'AC}$ は対角要素が0または1の対角行列でベキ等である．このことから

$$\mathbf{C'AC} = \mathbf{C'ACC'AC} = \mathbf{C'A^2C} \Rightarrow \mathbf{A} = \mathbf{A}^2 \tag{3b.4.5}$$

となる．

x_j^2 の特性関数は $\mu_j \neq 0$ ならば非心カイ2乗分布の特性関数になり (3b.4.3) と (3b.4.4) に余分な項が入る．読者はこのとき (3b.4.5) が成り立つことを導いてみよ．

カイ2乗分布の自由度は，rank $\mathbf{A}$，すなわち $\mathbf{A}$ がベキ等だから trace $\mathbf{A}$ に等しい [(i), 1b.7]．

(iii) <u>$\mathbf{Y'Y} = Q_1 + Q_2$，ただし $Q_1 \sim \chi^2(a)$ とする．このとき $Q_2 \sim \chi^2(n-a)$ である．</u>

$Q_1 = \mathbf{Y'AY}$ ならば (ii) より $\mathbf{A}^2 = \mathbf{A}$ である．一方 $Q_2 = \mathbf{Y'}(\mathbf{I}-\mathbf{A})\mathbf{Y}$ である．

$$(\mathbf{I}-\mathbf{A})^2 = \mathbf{I} + \mathbf{A}^2 - 2\mathbf{A} = \mathbf{I} - \mathbf{A}$$

だから $\mathbf{I}-\mathbf{A}$ はベキ等である．したがって上の結果が得られる．

(iv) <u>$Q = Q_1 + Q_2$．ただし $Q \sim \chi^2(a)$，$Q_1 \sim \chi^2(b)$，Q_2 は非負である．このとき $Q_2 \sim \chi^2(a-b)$ である．</u>

$Q \sim \chi^2(a)$ だから (3b.4.2) により $Q \to x_1^2 + \cdots + x_a^2$，$\mathbf{Y'Y} \to x_1^2 + \cdots + x_n^2$ となる直交変換が存在する．$Q_1 \to \mathbf{X'B_1X}$，$Q_2 \to \mathbf{X'B_2X}$ とし，ともに非負とする．

$$x_1^2 + \cdots + x_a^2 = \mathbf{X'B_1X} + \mathbf{X'B_2X}$$

だから右辺のそれぞれの2次形式は $x_1, \cdots, x_a$ のみを含まなければならない．このとき (iii) の結果が適用でき $Q_2 \sim \chi^2(a-b)$ が証明できる．

この結果から，もし $\mathbf{A}, \mathbf{B}, \mathbf{A}-\mathbf{B}$ が非負の2次形式の行列で，$\mathbf{A}$ と $\mathbf{B}$ がベキ等ならば $\mathbf{A}-\mathbf{B}$ もベキ等である．この結果を直接証明すれば (iv) の結果がでる．

(v) <u>$\mathbf{Y'A_1Y} \sim \chi^2(a)$，$\mathbf{Y'A_2Y} \sim \chi^2(b)$ とする．これらが互いに独立な分布に従うための必要十分条件は $\mathbf{A_1A_2} = \mathbf{0}$ である．</u>

$\mathbf{A}_1^2 = \mathbf{A}_1$，$\mathbf{A}_2^2 = \mathbf{A}_2$ だから $\mathbf{A_1A_2} = \mathbf{0}$ ならば

$$\mathbf{A}_1(\mathbf{I}-\mathbf{A}_1-\mathbf{A}_2) = \mathbf{A}_2(\mathbf{I}-\mathbf{A}_1-\mathbf{A}_2) = \mathbf{0}$$

である．これは

$$\operatorname{rank}(\mathbf{A}_1) + \operatorname{rank}(\mathbf{A}_2) + \operatorname{rank}(\mathbf{I}-\mathbf{A}_1-\mathbf{A}_2) = n$$

を意味している．一方 $\mathbf{Y'Y} = \mathbf{Y'A_1Y} + \mathbf{Y'A_2Y} + \mathbf{Y'}(\mathbf{I}-\mathbf{A}_1-\mathbf{A}_2)\mathbf{Y}$ である．したがってフィッシャー＝コクランの定理を適用すると2次形式は互いに独立な分布に従うことが示される．必要性は $\mathbf{Y'}(\mathbf{A}_1+\mathbf{A}_2)\mathbf{Y} \sim \chi^2(a+b)$ だから $\mathbf{A}_1+\mathbf{A}_2$ がベキ等である条件を示せば容易に証明できる．

(vi) $\mathbf{Y}'\mathbf{Y} = \mathbf{Y}'\mathbf{A}_1\mathbf{Y} + \cdots + \mathbf{Y}'\mathbf{A}_k\mathbf{Y}$ とする．このとき次のどちらの条件も $\mathbf{Y}'\mathbf{A}_i\mathbf{Y} \sim \chi^2(n_i, \lambda_i)$ であり，かつ集合 $\mathbf{Y}'\mathbf{A}_1\mathbf{Y}, \cdots, \mathbf{Y}'\mathbf{A}_k\mathbf{Y}$ が独立な分布に従うための必要十分条件である．ただし n_i は $\mathbf{A}_i$, $i = 1, \cdots, k$ の階数である．

(a) $\mathbf{A}_1, \cdots, \mathbf{A}_k$ の各々がベキ等行列．

(b) すべての $i \neq j$ にたいして $\mathbf{A}_i\mathbf{A}_j = 0$.

$\mathbf{A}_i^2 = \mathbf{A}_i$ だから $\operatorname{trace} \mathbf{A}_i = \operatorname{rank} \mathbf{A}_i$, $\mathbf{I} = \mathbf{A}_1 + \cdots + \mathbf{A}_k \Rightarrow \operatorname{trace} \mathbf{I} = n = \sum \operatorname{trace} \mathbf{A}_i = \sum n_i$. したがってフィッシャー＝コクランの定理が適用でき，必要十分条件 (a) が証明された．次に $\mathbf{I} = \mathbf{A}_1 + \cdots + \mathbf{A}_k$, $\mathbf{A}_i\mathbf{A}_j = 0 \Rightarrow \mathbf{A}_i(\mathbf{I} - \mathbf{A}_i) = 0$, すなわち $\mathbf{A}_i$ はベキ等である．したがって条件 (b) は条件 (a) と同等である．

上記の結果を $y_1, \cdots, y_n$ が相関をもつ場合，より正確にいうと (3b.3.5) の特性関数で定義される $\mathbf{Y} \sim N_n(\nu, \Sigma)$ の場合に拡張しよう．$R(\Sigma) = r$ のとき，ベクトル $\mathbf{Y}$ は

$$\mathbf{Y} = \nu + \mathbf{B}\mathbf{X} \tag{3b.4.6}$$

と書ける．ただし $\mathbf{X}$ は独立な $N(0, 1)$ の確率変数からなる r 次元ベクトルで $\mathbf{B}\mathbf{B}' = \Sigma$ である．$\mathbf{Y}$ の2次形式の研究は (3b.4.6) の関係により $\mathbf{X}$ の，すなわち独立な正規確率変数の2次関数の研究に帰着される．

(vii) Ogasawara and Takahashi(1951)　$\mathbf{Y} \sim N_n(\nu, \Sigma)$ とする．$(\mathbf{Y} - \nu)'\mathbf{A}(\mathbf{Y} - \nu)$ がカイ2乗分布に従うための必要十分条件は

$$\Sigma\mathbf{A}\Sigma\mathbf{A}\Sigma = \Sigma\mathbf{A}\Sigma \tag{3b.4.7}$$

である．このとき自由度は $\operatorname{tr}(\mathbf{A}\Sigma)$ である．

(3b.4.6) を用いると $(\mathbf{Y} - \nu)'\mathbf{A}(\mathbf{Y} - \nu) = \mathbf{X}'\mathbf{B}'\mathbf{A}\mathbf{B}\mathbf{X}$ である．次に (ii) の結果を適用すれば，必要十分条件は $\mathbf{B}'\mathbf{A}\mathbf{B}$ がベキ等，

$$\mathbf{B}'\mathbf{A}\mathbf{B}\mathbf{B}'\mathbf{A}\mathbf{B} = \mathbf{B}'\mathbf{A}\mathbf{B} \Leftrightarrow \mathbf{B}\mathbf{B}'\mathbf{A}\mathbf{B}\mathbf{B}'\mathbf{A}\mathbf{B}\mathbf{B}' = \mathbf{B}\mathbf{B}'\mathbf{A}\mathbf{B}\mathbf{B}'$$

すなわち $\Sigma\mathbf{A}\Sigma\mathbf{A}\Sigma = \Sigma\mathbf{A}\Sigma$ である．自由度は $\operatorname{tr} \mathbf{B}'\mathbf{A}\mathbf{B} = \operatorname{tr} \mathbf{A}\mathbf{B}\mathbf{B}' = \operatorname{tr} \mathbf{A}\Sigma$ である．

もし $|\Sigma| \neq 0$ ならば条件 (3b.4.7) は

$$\mathbf{A}\Sigma\mathbf{A} = \mathbf{A} \tag{3b.4.8}$$

となる．

(viii) $\mathbf{Y} \sim N_n(\nu, \Sigma)$ とする．$\mathbf{P}'\mathbf{Y}$ と $(\mathbf{Y} - \nu)'\mathbf{A}(\mathbf{Y} - \nu)$ が独立な分布に従うための必要十分条件は

$$\Sigma\mathbf{A}\Sigma\mathbf{P} = 0 \tag{3b.4.9}$$

である．

(ix) $\mathbf{Y} \sim N_n(\nu, \Sigma)$ とする．$(\mathbf{Y} - \nu)'\mathbf{A}(\mathbf{Y} - \nu)$ と $(\mathbf{Y} - \nu)'\mathbf{B}(\mathbf{Y} - \nu)$ が独立な分布に従うための必要十分条件は

$$\Sigma\mathbf{A}\Sigma\mathbf{B}\Sigma = 0 \tag{3b.4.10}$$

である．

(viii) と (ix) の結果は独立な確率変数の場合と同様な方法で証明できる．

その他の結果については Rao and Mitra (1971) の第9章を参照せよ．

2次形式の分布についてはたくさんの文献がある．3b.4 の結果は次のいくつかの文献に基づいている．Cochran(1934), Craig(1943), Hogg and Craig(1958), Ogasawara and Takahashi(1951), Ogawa(1949), Rao(1951a, 1953b, 1961a), Sakamoto(1944). さらに詳しい文献については Graybill (1961) を見よ．

3b.5 最小2乗法の3つの基本定理

独立な確率変数

$$y_i \sim N(x_{i1}\beta_1 + \cdots + x_{im}\beta_m, \sigma^2), \quad i = 1, \cdots, n \tag{3b.5.1}$$

を考えよう．x_{ij} は既知の係数，β_i は未知の母数である．行列の記号で $\mathbf{Y}$ が確率変数 y_i の列ベクトルをあらわし，β が母数 β_i の列ベクトルをあらわし，$\mathbf{X} = (x_{ij})$ が係数の行列をあらわすとすると

$$\sum (y_i - x_{i1}\beta_1 - \cdots - x_{im}\beta_m)^2 = (\mathbf{Y} - \mathbf{X}\beta)'(\mathbf{Y} - \mathbf{X}\beta)$$

となる．したがって $y_1, \cdots, y_n$ の密度関数は

$$c \cdot e^{-(\mathbf{Y}-\mathbf{X}\beta)'(\mathbf{Y}-\mathbf{X}\beta)/2\sigma^2} \tag{3b.5.2}$$

と書ける．この節では（第4章で論じる）最小2乗法の基礎となる，ある統計量の分布を導こう．

(i) **第1基本定理**

$$R_0^2 = \min_{\beta}(\mathbf{Y} - \mathbf{Y}\beta)'(\mathbf{Y} - \mathbf{X}\beta)$$

とする．このとき $R_0^2 \sim \sigma^2\chi^2(n - r)$ である．ただし r は $\mathbf{X}$ の階数である．

$\mathbf{F}$ をその列ベクトルが $\mathscr{M}(\mathbf{X})$ の正規直交基底となるような $n \times r$ 型行列とする．$(\mathbf{F} \vdots \mathbf{G})$ が直交行列になるように $n \times (n - r)$ 行列 $\mathbf{G}$ を行列 $\mathbf{F}$ に追加する．このような行列をつくると $\mathbf{F}'\mathbf{G} = 0$, $\mathbf{X}'\mathbf{G} = 0$, $\mathbf{G}'\mathbf{X} = 0$ となり，$\mathbf{Z}_1 = \mathbf{F}'\mathbf{Y}$, $\mathbf{Z}_2 = \mathbf{G}'\mathbf{Y}$ となるような $\mathbf{Y}$ から

$$\mathbf{Z} = \begin{pmatrix} \mathbf{Z}_1 \\ \mathbf{Z}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{F}' \\ \mathbf{G}' \end{pmatrix}\mathbf{Y}$$

への変換は直交変換である．ベクトル $\mathbf{X}\beta$ は，$\zeta_1 = \mathbf{F}'\mathbf{X}\beta$, $\zeta_2 = \mathbf{G}'\mathbf{X}\beta = 0$ に変換される．したがって (3b.1.2)（直交変換の下での距離の不変性）を適用すれば

$$(\mathbf{Y} - \mathbf{X}\beta)'(\mathbf{Y} - \mathbf{X}\beta) = (\mathbf{Z}_1 - \zeta_1)'(\mathbf{Z}_1 - \zeta_1) + \mathbf{Z}_2'\mathbf{Z}_2$$

となる．ゆえに，$\mathbf{Z}$ の密度関数は

$$c\cdot e^{-[(\mathbf{Z}_1-\mathbf{F}'\mathbf{X}\beta)'(\mathbf{Z}_1-\mathbf{F}'\mathbf{X}\beta)+\mathbf{Z}_2'\mathbf{Z}_2]/2\sigma^2}$$

である．これは $\mathbf{Z}_1$ と $\mathbf{Z}_2$ が互いに独立な分布に従うことを示している．$\mathbf{Z}_2$ のみの密度関数は

$$c\cdot e^{-\mathbf{Z}_2'\mathbf{Z}_2/2\sigma^2} \tag{3b.5.3}$$

である．一方，$\mathbf{Z}_2$ は $n-r$ 個の確率変数のベクトルで，この確率変数は (3b.5.3) によりすべて独立で $N(0,\sigma^2)$ に従う．よって $\mathbf{Z}_2'\mathbf{Z}_2\sim\sigma^2\chi^2(n-r)$ である．さて $\mathbf{Z}_1=\mathbf{F}'\mathbf{X}\beta$ となるような β を選ぶならば

$$\begin{aligned}R_0^2&=\min_\beta(\mathbf{Y}-\mathbf{X}\beta)'(\mathbf{Y}-\mathbf{X}\beta)=\min_\beta[(\mathbf{Z}_1-\zeta_1)'(\mathbf{Z}_1-\zeta_1)+\mathbf{Z}_2'\mathbf{Z}_2]\\&=\left[\min_\beta(\mathbf{Z}_1-\mathbf{F}'\mathbf{X}\beta)'(\mathbf{Z}_1-\mathbf{F}'\mathbf{X}\beta)\right]+\mathbf{Z}_2'\mathbf{Z}_2=\mathbf{Z}_2'\mathbf{Z}_2\end{aligned} \tag{3b.5.4}$$

となる．下に示すように方程式の数 r が $\mathbf{F}'\mathbf{X}$ の階数に等しいから，このような β の選択は可能である（8ページ条件4を参照）．

$$r=\operatorname{rank}\mathbf{X}'=\operatorname{rank}\mathbf{X}'(\mathbf{F}\,\vdots\,\mathbf{G})=\operatorname{rank}(\mathbf{X}'\mathbf{F}\,\vdots\,\mathbf{X}'\mathbf{G})=\operatorname{rank}(\mathbf{X}'\mathbf{F}\,\vdots\,0)=\operatorname{rank}\mathbf{X}'\mathbf{F}.$$

注意　R_0^2 の分布の問題にたいする射影作用素の知識を用いる，より簡潔な（別の）接近法は次の通りである．

$(\mathbf{Y}-\mathbf{X}\beta)'(\mathbf{Y}-\mathbf{X}\beta)$ は $\mathbf{X}\beta$ が $\mathbf{Y}$ の $\mathscr{M}(\mathbf{X})$ への射影であるとき最小となる．一方任意のベクトルの $\mathscr{M}(\mathbf{X})$ への射影は作用素，つまり対称でかつベキ等で，その階数が $\mathbf{X}$ の階数 r に等しい行列 $\mathbf{P}$ により得られる（1c.4 参照）．このように $\mathscr{M}(\mathbf{X})$ への $\mathbf{Y}$ の射影は $\mathbf{PY}$ となり，したがって $\mathbf{Y}-\mathbf{PY}$ は垂線となる．これより

$$R_0^2=(\mathbf{Y}-\mathbf{PY})'(\mathbf{Y}-\mathbf{PY})=\mathbf{Y}'(\mathbf{I}-\mathbf{P})(\mathbf{I}-\mathbf{P})\mathbf{Y}=\mathbf{Y}'(\mathbf{I}-\mathbf{P})\mathbf{Y}$$

となる．行列 $\mathbf{I}-\mathbf{P}$ もまたベキ等だから

$$\operatorname{rank}(\mathbf{I}-\mathbf{P})=\operatorname{trace}(\mathbf{I}-\mathbf{P})=\operatorname{trace}\mathbf{I}-\operatorname{trace}\mathbf{P}=n-r$$

である．そこで2次形式の分布についての［(ii)，3b.4］の必要十分条件を適用すると，

$$\mathbf{Y}'(\mathbf{I}-\mathbf{P})\mathbf{Y}\sim\sigma^2\chi^2(n-r,\lambda) \tag{3b.5.5}$$

となる．ただし λ は2次形式 $\mathbf{Y}'(\mathbf{I}-\mathbf{P})\mathbf{Y}/\sigma^2$ に，$\mathbf{Y}$ の期待値を代入したときの値である（(3b.2.2) 参照）．このようにして $\mathbf{PX}=\mathbf{X}$（$\mathscr{M}(\mathbf{X})$ のベクトルは射影作用素 $\mathbf{P}$ によって変わらないから）だから

$$\begin{aligned}\sigma^2\lambda&=E(\mathbf{Y}')(\mathbf{I}-\mathbf{P})E(\mathbf{Y})\\&=\beta'\mathbf{X}'(\mathbf{I}-\mathbf{P})\mathbf{X}\beta=\beta'(\mathbf{X}'\mathbf{X}-\mathbf{X}'\mathbf{PX})\beta=0\end{aligned}$$

となる．したがって分布 (3b.5.5) は実は中心カイ2乗分布である．

（ii）第2基本定理　$\mathbf{H}$ を $\mathscr{M}(\mathbf{H})\subset\mathscr{M}(\mathbf{X}')$ となるような $m\times k$ で階数 k の行列とし，条件 $\mathbf{H}'\beta=\xi$（所与）の下で

$$R_1^2=\min_\beta(\mathbf{Y}-\mathbf{X}\beta)'(\mathbf{Y}-\mathbf{X}\beta)$$

とする．このとき

（a）R_0^2 と $R_1^2-R_0^2$ は互いに独立な分布に従う．

（b）$R_0^2\sim\sigma^2\chi^2(n-r)$，$R_1^2-R_0^2$ は自由度 k の非心カイ2乗分布に従う．

（c）$\mathbf{H}'\beta=\xi$ が成りたつならば $R_1^2-R_0^2\sim\sigma^2\chi^2(k)$，

$$\frac{R_1^2-R_0^2}{k}\div\frac{R_0^2}{n-r}\sim F(k,n-r)$$

である．

$\mathscr{M}(\mathbf{X}')=\mathscr{M}(\mathbf{X}'\mathbf{X})$ だから条件 $\mathscr{M}(\mathbf{H})\subset\mathscr{M}(\mathbf{X}')\Rightarrow\mathscr{M}(\mathbf{H})\subset\mathscr{M}(\mathbf{X}'\mathbf{X})$ となる．したがって (3b.1.4) により $\mathbf{H}=\mathbf{X}'\mathbf{XC}$，$\operatorname{rank}\mathbf{XC}=k$ となるような行列 $\mathbf{C}$ が存在する．

$$\mathbf{Z}_1=\mathbf{D}'\mathbf{Y},\quad \mathbf{Z}_2=\mathbf{G}'\mathbf{Y},\quad \mathbf{Z}_3=\mathbf{C}'\mathbf{X}'\mathbf{Y}$$

なる変換を行う．ただし $\mathbf{G}$ は (i) での行列，$\mathbf{D}'$ は $\mathbf{D}'\mathbf{D}=\mathbf{I}$ となるように選ばれ，$\mathbf{G}$ と $\mathbf{XC}$ に直交している．$\mathbf{Z}_1,\mathbf{Z}_2,\mathbf{Z}_3$ の確率変数の個数は，それぞれ，$r-k,n-r,k$ である．ベクトル $\mathbf{X}\beta$ は

$$\zeta_1=\mathbf{D}'\mathbf{X}\beta,\quad \zeta_2=\mathbf{G}'\mathbf{X}\beta=0,\quad \zeta_3=\mathbf{C}'\mathbf{X}'\mathbf{X}\beta=\mathbf{H}'\beta$$

に変換される．変換の部分行列は互いに直交するから (3b.1.3) の結果が適用できる．したがって

$$\begin{aligned}(\mathbf{Y}-\mathbf{X}\beta)'(\mathbf{Y}-\mathbf{X}\beta)&=(\mathbf{Z}_1-\mathbf{D}'\mathbf{X}\beta)'(\mathbf{Z}_1-\mathbf{D}'\mathbf{X}\beta)+\mathbf{Z}_2'\mathbf{Z}_2\\&\quad+(\mathbf{Z}_3-\mathbf{H}'\beta)'(\mathbf{C}'\mathbf{X}'\mathbf{XC})^{-1}(\mathbf{Z}_3-\mathbf{H}'\beta)\end{aligned} \tag{3b.5.6}$$

となる．$\mathbf{Z}_1,\mathbf{Z}_2,\mathbf{Z}_3$ の密度関数を書くと，それらが独立な分布に従っていることがわかる．特に $\mathbf{Z}_3$ の密度関数は

$$c\cdot e^{-(\mathbf{Z}_3-\mathbf{H}'\beta)'(\mathbf{C}'\mathbf{X}'\mathbf{XC})^{-1}(\mathbf{Z}_3-\mathbf{H}'\beta)/2\sigma^2}$$

であり，これは正規確率変数の k 個の1次関数の密度関数と同じである．

(3b.4.7) の結果を用いると，2次形式

$$Q=(\mathbf{Z}_3-\xi)'(\mathbf{C}'\mathbf{X}'\mathbf{XC})^{-1}(\mathbf{Z}_3-\xi)\sim\sigma^2\chi^2(k,\lambda)$$

となる．ここで非心率は (3b.2.2) より $\lambda=\sigma^{-2}(\mathbf{H}'\beta-\xi)'(\mathbf{C}'\mathbf{X}'\mathbf{XC})^{-1}(\mathbf{H}'\beta-\xi)$ である．

さらに (i) で示したように $R_0^2=\mathbf{Z}_2'\mathbf{Z}_2\sim\sigma^2\chi^2(n-r)$ で R_0^2 は Q とは独立である．$\mathbf{H}'\beta=\xi$ が成り立つとき $\lambda=0$ である．この場合 $Q\sim\sigma^2\chi^2(k)$ となる．結果 (a)，(b) そして (c) は $Q=R_1^2-R_0^2$ を示せば証明できるだろう．

$\mathbf{Z}_1-\mathbf{D}'\mathbf{X}\beta=0$，$\mathbf{H}'\beta=\xi$ となるような β が存在するならば

$$\begin{aligned}R_1^2&=\min_{\mathbf{H}'\beta=\xi}(\mathbf{Y}-\mathbf{X}\beta)'(\mathbf{Y}-\mathbf{X}\beta)\\&=[\min_{\mathbf{H}'\beta=\xi}(\mathbf{Z}_1-\mathbf{D}'\mathbf{X}\beta)'(\mathbf{Z}_1-\mathbf{D}'\mathbf{X}\beta)]+R_0^2+Q\end{aligned}$$

［(3b.5.6) を用いて］

$$= R_0^2 + Q$$

となる．方程式の数が r で

$$r = \operatorname{rank} \mathbf{X}' = \operatorname{rank} \mathbf{X}'(\mathbf{D} \vdots \mathbf{G} \vdots \mathbf{XC}) = \operatorname{rank}(\mathbf{X}'\mathbf{D} \vdots 0 \vdots \mathbf{H}) = \operatorname{rank}(\mathbf{X}'\mathbf{D} \vdots \mathbf{H})$$

だからこのような β は存在する．

注意　射影作用素を用いる別の接近法は次の通りである．まず $\mathbf{H}'\beta = \xi$ ならば $\beta = \beta_0 + \gamma$ である．ただし β_0 は $\mathbf{H}'\beta = \xi$ の特殊解であり，γ は $\mathbf{H}'\beta = 0$ の一般解である．したがって

$$\min_{\mathbf{H}'\beta=\xi}(\mathbf{Y} - \mathbf{X}\beta)'(\mathbf{Y} - \mathbf{X}\beta) = \min_{\mathbf{H}'\gamma=0}(\mathbf{Y} - \mathbf{X}\beta_0 - \mathbf{X}\gamma)'(\mathbf{Y} - \mathbf{X}\beta_0 - \mathbf{X}\gamma) \tag{3b.5.7}$$

である．一方 $\mathbf{H}'\gamma = 0$ なる制約をもつ $\mathbf{X}\gamma$ は，$\mathbf{H}$ が条件 $\mathscr{M}(\mathbf{H}) \subset \mathscr{M}(\mathbf{X}')$ をみたすかどうかにかかわらず $d[\mathscr{S}] = \operatorname{rank}(\mathbf{X}' \vdots \mathbf{H}) - \operatorname{rank} \mathbf{H} = s$（とおく）の部分空間 $\mathscr{S} \subset \mathscr{M}(\mathbf{X})$ である．$\mathbf{P}$ を，ベクトルを $\mathscr{M}(\mathbf{X})$ へ射影する作用素とし，$\mathscr{S}$ へ射影する作用素を $\mathbf{U}$ とする．すると $\operatorname{rank} \mathbf{P} = r$，$\operatorname{rank} \mathbf{U} = s$ である．次に

$$R_1^2 = (\mathbf{Y} - \mathbf{X}\beta_0)'(\mathbf{I} - \mathbf{U})(\mathbf{Y} - \mathbf{X}\beta_0)$$

$$R_0^2 = \mathbf{Y}'(\mathbf{I} - \mathbf{P})\mathbf{Y} = (\mathbf{Y} - \mathbf{X}\beta_0)'(\mathbf{I} - \mathbf{P})(\mathbf{Y} - \mathbf{X}\beta_0).$$

R_0^2 の式に因子 $\mathbf{X}\beta_0$ を導入しても R_0^2 の値は変らない．$\mathbf{I} - \mathbf{U}$ はベキ等で階数が $n - s$ だから $R_1^2 \sim \sigma^2\chi^2(n - s, \lambda)$ である．ただし

$$\sigma^2\lambda = (\mathbf{X}\beta - \mathbf{X}\beta_0)'(\mathbf{I} - \mathbf{U})(\mathbf{X}\beta - \mathbf{X}\beta_0) \tag{3b.5.8}$$

である．さて第1基本定理により $R_0^2 \sim \sigma^2\chi^2(n - r)$ で，$R_1^2 - R_0^2 \geqslant 0$ である．ここで [(iv), 3b.4] を適用すると $R_1^2 - R_0^2 \sim \sigma^2\chi^2(r - s, \lambda)$ となり，R_0^2 とは独立である．

$\mathbf{H}'\beta = \xi$ が成り立つならば $\beta = \beta_0 + \gamma$ である．ただし $\mathbf{H}'\gamma = 0$ である．したがって $\mathbf{UX}\gamma = \mathbf{X}\gamma$ だから

$$\sigma^2\lambda = (\mathbf{X}\gamma)'(\mathbf{I} - \mathbf{U})(\mathbf{X}\gamma) = (\mathbf{X}\gamma)'(\mathbf{X}\gamma) - (\mathbf{X}\gamma)'(\mathbf{X}\gamma) = 0$$

となる．このようにして $R_1^2 - R_0^2 \sim \sigma^2\chi^2(r - s)$，つまり中心カイ2乗分布に従う．

$\mathscr{M}(\mathbf{H}) \subset \mathscr{M}(\mathbf{X}')$ で $\mathbf{H}$ の階数が k であるような特別な場合では

$$s = \operatorname{rank}(\mathbf{X}' \vdots \mathbf{H}) - \operatorname{rank} \mathbf{H} = r - k \Rightarrow r - s = k \tag{3b.5.9}$$

となり，$R_1^2 - R_0^2 \sim \sigma^2\chi^2(k)$ である．

（iii）第3基本定理　$\mathbf{Y}' = (\mathbf{Y}_1' \vdots \mathbf{Y}_2')$ とし，これに対応する期待値のベクトルの分割を $\beta'\mathbf{X}' = \beta'(\mathbf{X}_1' \vdots \mathbf{X}_2') = (\beta'\mathbf{X}_1' \vdots \beta'\mathbf{X}_2')$ とする．このとき統計量

$$U = \frac{\min_\beta(\mathbf{Y}_1 - \mathbf{X}_1\beta)'(\mathbf{Y}_1 - \mathbf{X}_1\beta)}{\min_\beta(\mathbf{Y} - \mathbf{X}\beta)'(\mathbf{Y} - \mathbf{X}\beta)} \sim B\left(\frac{q - r_1}{2}, \frac{n - q - r + r_1}{2}\right) \tag{3b.5.10}$$

である．ただし q は $\mathbf{X}_1'$ の列の数，$r_1 = \operatorname{rank} \mathbf{X}_1$，$n, r$ は（i）で定義した通りである．

（ii）と同様な変換を行えば統計量 U は

$$\frac{\chi^2(q - r_1)}{\chi^2(q - r_1) + \chi^2(n - r - q + r_1)}$$

であらわすことができる．ここで2つのカイ2乗確率変数は独立である．したがって(3b.1.6) よりこの比はベータ分布 (3b.5.10) に従う．

（iv）一般定理　$\mathbf{W}$ を p 次元ベクトルとし，$\mathbf{Y}$ を与えたとき $\mathbf{W}$ の条件付分布が

$$\mathbf{W} \mid \mathbf{Y} \sim N_p(\mathbf{Y}\alpha, \sigma^2\mathbf{G}),$$

$$R(\mathbf{Y}'\mathbf{G}\mathbf{Y}) = r \quad (\text{一定}) \qquad (\text{確率1で})$$

$$R(\mathbf{G}) = s, \qquad (\mathbf{G} \text{ は } \mathbf{Y} \text{ によらない}) \tag{3b.5.11}$$

となるような，確率変数の $p \times k$ 行列を $\mathbf{Y}$ とする．任意に選んだ一般逆行列にたいし次の統計量を定義する．

$$T_1 = \mathbf{d}'\mathbf{Y}(\mathbf{Y}'\mathbf{G}\mathbf{Y})^-\mathbf{Y}'\mathbf{d},$$

$$T_2 = \mathbf{d}'\mathbf{G}^-\mathbf{d}, \qquad \mathbf{d} = \mathbf{W} - \mathbf{Y}\alpha. \tag{3b.5.12}$$

このとき $T_1 \sim \sigma^2\chi^2(r)$，$T_2 - T_1 \sim \sigma^2\chi^2(s - r)$ で互いに独立な分布に従い，

$$\frac{(s - r)T_1}{r(T_2 - T_1)}$$

は $F(r, s - r)$ 分布に従う．

この結果は 3b.4 の命題 (iv)～(ix) をいいかえたものである．もちろん $\mathbf{Y}$ が確率変数でないときもこの定理は成り立つ．

別の証明法や，さらに進んだ結果についてはこの章の終わりの問題2.1—2.11 を見よ．

3b.6　p 変量正規分布

(3b.3.4) に従って密度関数が

$$c \cdot e^{-(\mathbf{X}-\mu)'\Sigma^{-1}(\mathbf{X}-\mu)/2} \tag{3b.6.1}$$

である p 変量正規分布を考えよう．ここで正定符号行列 Σ，p 次元ベクトル μ は分布の母数である．

ルベーグ測度に関して $\mathbf{X}$ の密度が存在しないような場合も含めた，より一般的な定義は第8章で与える．しかし (3a.6.1) は古典的な興味をもたせ，その性質を調べることは興味のあることであろう．

（i）(3b.6.1) の定数 c の値は

$$(2\pi)^{-p/2}|\Sigma|^{-1/2} \tag{3b.6.2}$$

である．

$\mathbf{Y} = \Sigma^{-1/2}\mathbf{X}$ なる変換を行う．ただし $\Sigma^{1/2}$ は Σ の平方根である．微分 $d\mathbf{Y} = |\Sigma|^{-1/2}d\mathbf{X}$ だから $\mathbf{Y}$ の密度関数は

$$c\cdot|\Sigma|^{1/2}e^{-(\mathbf{Y}-\nu)'(\mathbf{Y}-\nu)/2},\qquad \nu=\Sigma^{-1/2}\mu \tag{3b.6.3}$$

である．$\mathbf{Y}$ の成分は互いに独立で正規分布に従う．よって(3b.6.3) を項別積分すると $c\cdot|\Sigma|^{1/2}(2\pi)^{p/2}=1$ が得られ，これから (3b.6.2) の c の値がでる．

変換 $\mathbf{Y}=\Sigma^{-1/2}\mathbf{X}$ の意義に注意しよう．それは，p 変量正規分布の研究が **3b.2—3b.5** で考えた p 個の独立な1変量正規確率変数の研究に帰着することを示している．たとえば p 変量正規確率変数 $\mathbf{X}$ に基づく統計量 $f(\mathbf{X})$ の性質をしらべるには独立な1変量正規確率変数 $\mathbf{Y}$ に基づく統計量 $f(\Sigma^{1/2}\mathbf{Y})$ を考えさえすればよい．以下の命題でこのことを説明しよう．

（ii） $E(\mathbf{X})=\mu$，すなわち μ は $\mathbf{X}$ の成分の平均値のベクトルである．

$\mathbf{X}=\Sigma^{1/2}\mathbf{Y}$ だから (3b.6.3) を用いると $E(\mathbf{X})=\Sigma^{1/2}E(\mathbf{Y})=\Sigma^{1/2}\Sigma^{-1/2}\mu=\mu$ となる．

（iii） $D(\mathbf{X})$ つまり $\mathbf{X}$ の分散共分散行列は Σ である．

再び変換 $(\mathbf{X}-\mu)=\Sigma^{1/2}(\mathbf{Y}-\nu)$ を用いると $(\mathbf{Y}-\nu)$ の成分は独立な $N(0,1)$ に従う確率変数だから $D(\mathbf{Y})=\mathbf{I}$ である．このとき

$$\begin{aligned}D(\mathbf{X})&=D[\Sigma^{1/2}(\mathbf{Y}-\nu)]=\Sigma^{1/2}D(\mathbf{Y}-\nu)\Sigma^{1/2}\\&=\Sigma^{1/2}\Sigma^{1/2}=\Sigma\end{aligned}$$

となる．

命題 (ii) と (iii) は p 変量正規分布が確率変数ベクトル $\mathbf{X}$ の平均と分散共分散で完全に規定されることをあらわしている．密度関数 (3b.6.1) をもつ分布を $N_p(\mu,\Sigma)$ と書く．

（iv） $\mathbf{X}$ の成分の任意の周辺分布，たとえば最初の k 個の成分の周辺分布は k 変量正規分布である．

$\mathbf{X}'=(\mathbf{X}_1'\,\vdots\,\mathbf{X}_2')$ とし，$\mathbf{X}_1$ は最初の k 個の成分からなる部分ベクトルとする．すると

$$\mathbf{X}'=(\mathbf{X}_1'\,\vdots\,\mathbf{X}_2')=\mathbf{Y}'\Sigma^{1/2}=\mathbf{Y}'(\mathbf{C}_1\,\vdots\,\mathbf{C}_2)\Rightarrow\mathbf{X}_1=\mathbf{C}_1'\mathbf{Y}$$

である．したがって (3b.3.4) を用いると $\mathbf{X}_1$ は k 変量正規分布，$N_k(\mu_1,\Sigma_{11})$ に従う．ただし $E(\mathbf{X}_1)=\mu_1$，$D(\mathbf{X}_1)=\Sigma_{11}$ で μ_1 と Σ_{11} はそれぞれ μ と Σ を適当に分割したものである．

多変量正規分布のさらに進んだ研究については第8章を参照せよ．

3b.7　指数型分布族

$(\Omega,\mathscr{B},\mu)$ を測度空間とし，$T_1,\cdots,T_q$ は Ω 上の q 個の実数値可測関数をあらわすものとする．このとき典型的な指数型分布族は測度 μ に関して

$$f(\omega,\theta)=C(\theta)h(\omega)\exp[\textstyle\sum\pi_i(\theta)T_i(\omega)],\qquad \omega\in\Omega \tag{3b.7.1}$$

という形の確率密度によって特徴づけられる．ここで，指定した母数空間を Θ とすると $\theta\in\Theta$ であり π_i は既知の関数である．(3b.7.1) はまた

$$f(\omega,\theta)=\exp[\pi(\theta)+T(\omega)+\textstyle\sum\pi_i(\theta)T_i(\omega)] \tag{3b.7.2}$$

という別の形でも書くことができる．ここで $\pi(\theta)$ は離散的な場合は $\sum f(\omega,\theta)=1$，一般には $\int f(\omega,\theta)d\mu=1$ の関係から得られる．一般性を失なうことなく $1,T_1(\omega),\cdots,T_q(\omega)$ あるいは $1,\pi_1(\theta),\cdots,\pi_q(\theta)$ の間には線形関係がないと仮定できる．そうでなければ関数 $T_1,\cdots,T_q$ の間には関数関係があることになる．

多くの離散分布や連続分布は (3b.7.1) のクラスに属する．読者は2項分布，ポアソン分布，負の2項分布等の離散分布とガンマ分布，正規分布，一様分布等の連続分布がその例であることを確かめてみよ．

たとえば Ω が一般座標が $X_1,\cdots,X_p$ の空間 R^p ならば，**3b.6** で考えた p 変量正規分布の族は

$$(T_1,\cdots,T_q)=(X_1,\cdots,X_p,X_1^2,\cdots,X_p^2,X_1X_2,X_1X_3,\cdots,X_{p-1}X_p) \tag{3b.7.3}$$

できる．ただし $q=p+p(p+1)/2$ である．このように指数型分布族は特別な場合として多変量正規分布を含む．(3b.7.3) では $T_1,\cdots,T_q$ の間に関係があることに注意せよ．

密度関数の分解定理（**2d.3** 参照）を適用すると $(T_1,\cdots,T_q)$ は指数型分布族の十分統計量であることがわかる．指数型分布族の特別な場合として

$$f(\omega,\theta)=\exp[h(\theta)+T(\omega)+\theta_1T_1(\omega)+\cdots+\theta_qT_q(\omega)] \tag{3b.7.4}$$

がある．ここで $\theta=(\theta_1,\cdots,\theta_q)$ は集合 $\Theta\subset R^q$ に属する制約のない母数のベクトルと考えられる．このようなとき，以下の命題は興味がある．

（i） 自然な母数空間は凸である．

（ii） 統計量の部分集合 $(T_1,\cdots,T_r)$ の分布もまた指数型分布族に属する．

（iii） Θ が q 次元の長方形を含むならば集合 $(T_1,\cdots,T_q)$ は θ にたいして十分であり，その分布は完備である，すなわち

$$\text{各 }\theta\in\Theta\text{ にたいし }\quad Ef(T_1,\cdots,T_q)=0\Rightarrow f(T_1,\cdots,T_q)=0,\qquad \text{a. e.}$$

である．

モデル (3b.7.4) は空間 S で定義された変数の集合 $X_1,\cdots,X_p$ を導入することにより拡張できる．各応答 $\omega\in\Omega$ は応答に影響を与える観測可能なベクトル $(X_1,\cdots,X_p)$ と対になっている．拡張されたモデルは

$$f(\omega,\theta|X_1,\cdots,X_p)=\exp[\alpha+\textstyle\sum\sum\phi_{ij}X_iT_j(\omega)] \tag{3b.7.5}$$

で定義できる．ここで α は正規化定数で母数 ϕ_{ij} と X_i に依存する．モデル (3b. 7. 5)は特別な場合として多変量正規分布や Bock (1969) が考えた多項ロジットにおいて〔確率変数の部分集合を与えたときの残りの確率変数の〕条件付分布を含んでいる．このように (3b. 7. 1)，(3b. 7. 4)，(3b. 7. 5) で考えた指数型分布族は実際の研究で役に立つさまざまなモデルを与える (Dempster(1971) を参照)．

3c　対称正規分布

3c. 1　定　　義

確率変数 $x_1, \cdots, x_p$ の分布が平均 $E(x_i) = \mu$ (すべての i について同じ) で分散共分散行列が

$$\mathbf{D} = \sigma^2 \begin{pmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & \cdots & \rho \\ & \cdots\cdots\cdots\cdots & & \\ \rho & \rho & \cdots & 1 \end{pmatrix}$$

なる形をもつ p 変量正規分布ならば，$x_1, \cdots, x_p$ は対称正規分布に従うという．$\mathbf{C}$ の第1行が $(1/\sqrt{p}, \cdots, 1/\sqrt{p})$ である直交変換

$$\mathbf{Y} = \mathbf{CX} \tag{3c. 1. 1}$$

を行う．簡単な計算により

$$E(y_1) = \mu\sqrt{p}, \qquad V(y_1) = (1 + \overline{p-1}\rho)\sigma^2 \tag{3c. 1. 2}$$

$$E(y_i) = 0, \qquad V(y_i) = (1-\rho)\sigma^2, \qquad i = 2, \cdots, p \tag{3c. 1. 3}$$

$$\mathrm{cov}(y_i, y_j) = 0, \qquad i \neq j \text{ にたいして} \tag{3c. 1. 4}$$

すなわち変換された確率変数はすべて無相関であることがわかる．(3c. 1. 2)から(3c. 1. 4)までの結果は

$$y_1 \sim N(\mu\sqrt{p}, [1 + (p-1)\rho]\sigma^2)$$

$$y_i \sim N[0, (1-\rho)\sigma^2], \qquad i = 2, \cdots, p \tag{3c. 1. 5}$$

であり，かつ y_i はすべて独立であることを示している．$\mathbf{C}$ は正確には行列 $\mathbf{D}$ の固有ベクトルからなる行列である．

3c. 2　標　本　分　布

平均と分散の同時分布　$x_1, \cdots, x_p$ が対称正規分布に従うとする．このとき，平均 $\bar{x} = (x_1 + \cdots + x_p)/p$ と，分散 $W/(p-1) = \Sigma(x_i - \bar{x})^2 \div (p-1)$ は独立に分布し，かつ，次の関係がある．

$$\bar{x} \sim N(\mu, (1 + \overline{p-1}\rho)\sigma^2/p), \qquad W \sim (1-\rho)\sigma^2\chi^2(p-1)$$

直交変換 (3c. 1. 1) の下で

$$\sum x_i^2 = \sum y_i^2$$

$$y_1 = \sqrt{p}\,\bar{x} \quad \text{および} \quad W = y_2^2 + \cdots + y_p^2$$

が成り立ち，(3c. 1. 5) の結果により y_1 と W は独立な分布に従う．y_1 は (3c. 1. 5) のように正規分布に従い，$W \sim (1-\rho)\sigma^2\chi^2(p-1)$ である．

無相関な確率変数の場合 (3b. 3. 1) のように，観測値の標本平均と標本分散は独立な分布に従う．

3b. 3 のように t を次のように定義する．

$$t = \sqrt{p}(\bar{x} - \mu)/\sqrt{W/(p-1)}.$$

すると $t \sim \sqrt{\dfrac{1 + \overline{p-1}\rho}{1-\rho}} S(p-1)$ つまり $\sqrt{(1-\rho)}t/\sqrt{(1 + \overline{p-1}\rho)}$ が自由度 $p-1$ のスチューデントの分布，$S(p-1)$ に従うことが容易にわかる．しかし ρ が未知のときは μ についての推測はできない．

標本間と標本内の平方和　p 次元の対称正規確率変数の繰り返しのある観測値を考える．i 番目の観測値を $(x_{1i}, \cdots, x_{pi})$，$i = 1, \cdots, n$ とする．

$$\bar{x}_i = \frac{(x_{1i} + \cdots + x_{pi})}{p}$$

$$W_i = \sum_j (x_{ji} - \bar{x}_i)^2 \tag{3c. 2. 1}$$

$$B = p\sum(\bar{x}_i - \bar{x})^2, \qquad W = W_1 + \cdots + W_n$$

を定義する．ただし $\bar{x}$ は総平均である．$T = \sum\sum(x_{ij} - \bar{x})^2$ ならば $T = B + W$ の関係が得られる．B を標本間の平方和，W を標本内の平方和と呼ぶ．

各 i について $\bar{x}_i$ と W_i が独立であることはすでに証明されている．$W_1, \cdots, W_n$ は異なった標本から得られるので独立である．$W_i \sim (1-\rho)\sigma^2\chi^2(p-1)$ だから

$$W = W_1 + \cdots + W_n \sim (1-\rho)\sigma^2\chi^2(np-n) \tag{3c. 2. 2}$$

である．同様に $\bar{x}_1, \cdots, \bar{x}_n$ は独立で，それぞれ同じ平均と同じ分散 $(1 + \overline{p-1}\rho)\sigma^2/p$ の正規分布に従う．よって

$$\frac{B}{p} = \sum(\bar{x}_i - \bar{x})^2 \sim \frac{1 + \overline{p-1}\rho}{p}\sigma^2\chi^2(n-1)$$

すなわち

$$B \sim (1 + \overline{p-1}\rho)\sigma^2\chi^2(n-1) \tag{3c. 2. 3}$$

である．(3c. 2. 2) と (3c. 2. 3) の結果を結びつけ，(3b. 1. 10) を適用すると

$$\frac{B}{1 + \overline{p-1}\rho} \div \frac{W}{1-\rho} \sim f[n-1, n(p-1)] \tag{3c. 2. 4}$$

を得る．ここでfは2つのカイ2乗確率変数の比の分布である．$\rho=0$のとき

$$\frac{B}{W}\sim f[n-1,n(p-1)],\qquad \frac{W}{B}\sim f[n(p-1),n-1] \tag{3c.2.5}$$

となる．さらに (3b.1.12) を適用すると$\rho=0$のとき，

$$\frac{W}{W+B}=\frac{W}{T}\sim B\left(\frac{n(p-1)}{2},\frac{n-1}{2}\right) \tag{3c.2.6}$$

となる．

級内相関係数　n世帯のそれぞれについて，p人の兄弟の身長を測定し，それをもとに兄弟の身長の相関係数を推定しよう．1世帯でのp個の測定値から$p(p-1)$対の観測値(x,y)が得られる．ここでxはある兄弟の身長でyは他の兄弟の身長である．n世帯から合計$np(p-1)$対の観測値が得られ，それから通常の方法で相関係数を計算する．観測値を

$$(x_{1i},\cdots,x_{pi}),\qquad i=1,\cdots,n\ (\text{世帯}) \tag{3c.2.7}$$

とする．(3c.2.1) の記号を用いて (3c.2.7) から得られた$np(p-1)$対の(x,y)についての修正平方和Sと修正積和Pは$S=(p-1)T$，$P=pB-T$であり，したがって級内相関係数は

$$r=\frac{pB-T}{(p-1)T}$$

となり，これから

$$\frac{W}{T}=\frac{(1-r)(p-1)}{p}$$

の関係が成り立つことを示すのはよい練習問題であろう．$\rho=0$のとき (3c.2.6) から

$$\frac{W}{T}\sim B\left(\frac{n(p-1)}{2},\frac{n-1}{2}\right)$$

である．したがって

$$u=\frac{(1-r)(p-1)}{p}\sim B\left(\frac{n(p-1)}{2},\frac{n-1}{2}\right)$$

となる．uの密度関数は（ベータ分布を書けば）

$$c\cdot u^{[n(p-1)/2]-1}(1-u)^{[(n-1)/2]-1}$$

であり，r（級内相関係数）の密度関数は

$$c\cdot(1-r)^{[n(p-1)/2]-1}(1+r\overline{p-1})^{[(n-1)/2]-1} \tag{3c.2.8}$$

である．$\rho\neq 0$のとき，rの分布を求めるには，まず

$$\frac{B}{W}=\frac{1+(p-1)r}{(1-r)(p-1)}$$

だから (3c.2.4) を用いると

$$\frac{1+\overline{p-1}r}{(1-r)(p-1)}\cdot\frac{1-\rho}{1+(p-1)\rho}\sim f(n-1,n(p-1)) \tag{3c.2.9}$$

となる．次に (3c.2.9) から変数変換によりrの分布が得られる．

平均に関する対称性の検定　たとえばx_iがその世帯でi番目（順位i）の兄弟の身長をあらわすように，変数$(x_1,\cdots,x_p)$に自然な順位がついているとしよう．n世帯の観測値をもとにして，異なった順位の兄弟の平均身長に差があるかどうかを検定できるだろうか？

np個の観測値 (3c.2.7) からi番目の兄弟の平均身長

$$\bar{x}_{i\cdot}=\frac{(x_{i1}+\cdots+x_{in})}{n},\qquad i=1,\cdots,p$$

とその平方和

$$\sum(\bar{x}_{i\cdot}-\bar{x})^2$$

を計算しよう．$R=n\sum(\bar{x}_{i\cdot}-\bar{x})^2$とする．異なった順位間の差をあらわす統計量としてのRの分布を求めよう．

i番目の集合$(x_{1i},\cdots,x_{pi})$は (3c.1.1) の変換$\mathbf{Y}=\mathbf{CX}$により新しい変数

$$y_{1i},y_{2i},\cdots,y_{pi},\qquad i=1,\cdots,n$$

にうつることを思い出そう．y_{ij}はすべて独立で

$$V(y_{ij})=(1-\rho)\sigma^2,\qquad i>1,\ j=1,\cdots,n$$

である．次に

$$\sum_2^p\sum_1^n y_{ij}^2=n\sum_2^p\bar{y}_{i\cdot}^2+\sum_2^p\sum_1^n(y_{ij}-\bar{y}_{i\cdot})^2,$$

$$W=Q_1+Q_2$$

なる等式が成り立つ．$\sqrt{n}\,\bar{y}_{i\cdot}\sim N(0,(1-\rho)\sigma^2)$だから$Q_1\sim(1-\rho)\sigma^2\chi^2(p-1)$であり，$y_{ij}\sim N(0,(1-\rho)\sigma^2)$だから$W\sim(1-\rho)\sigma^2\chi^2[n(p-1)]$である．$Q_2$は明らかに正であり，[(iv), 3b.4] を適用すると

$$Q_2\sim(1-\rho)\sigma^2\chi^2[(n-1)(p-1)]$$

となりさらにQ_1とQ_2は互いに独立な分布に従う．

平均$\bar{x}_{1\cdot},\cdots,\bar{x}_{p\cdot}$は直交変換$\mathbf{Y}=\mathbf{CX}$により$\bar{y}_{1\cdot},\cdots,\bar{y}_{p\cdot}$にうつる．したがって

$$\sum\bar{x}_{i\cdot}^2=\sum\bar{y}_{i\cdot}^2$$

である．一方$\bar{x}$をすべてのnp個の観測値の平均とすると$\bar{y}_{1\cdot}=\sqrt{p}\,\bar{x}$である．したがって

$$\frac{R}{n}=\sum(\bar{x}_{i\cdot}-\bar{x})^2=\bar{y}_{2\cdot}^2+\cdots+\bar{y}_{p\cdot}^2=\frac{Q_1}{n}$$

となる．ゆえに，RはQ_1と同じ分布，すなわち$(1-\rho)\sigma^2\chi^2(p-1)$に従う．$R$の分布

は未知の母数を含むから，次の統計量を定義する．

$$F = \frac{R}{p-1} \div \frac{Q_2}{(n-1)(p-1)}. \tag{3c. 2. 10}$$

(3c. 2. 10) の統計量 F は未知の母数 ρ と σ に無関係なフィッシャーの分布 $F[p-1, (n-1)(p-1)]$ に従う．Q_2 が順位の差を含まないから統計量 F は R で測った順位による差をあらわしている．したがって検定の際，仮説にたいして適当な基準を設けることができる．すでに (3c. 2. 1) で定義した T, B は $T = B + W = B + R + Q_2$ をみたすことに注意しよう．このようにして Q_2 は式 $T - B - R$ から（もとの観測値 x_{ij} を使って）計算できる．T, B, R を簡単に計算できる形であらわすと

$$T = \sum\sum x_{ij}^2 - np\bar{x}^2, \qquad B = \frac{1}{p}\sum x_{\cdot j}^2 - np\bar{x}^2$$

$$R = \frac{1}{n}\sum x_{i\cdot}^2 - np\bar{x}^2$$

である．ただし $x_{\cdot j} = \sum_i x_{ij}$, $x_{i\cdot} = \sum_j x_{ij}$ である．

3d　2変量正規分布

3d. 1　一般的性質

多変量正規分布は，独立な正規確率変数の1次関数の標本分布を与えるものとしてすでに導入した．2変量正規分布は，むろんその特別な場合であるが，その性質を調べることはいくらか興味がある．一般の多変量正規分布は第8章で，異なった観点からくわしく論じる．

1対の変数 (x, y) はその密度関数が

$$c \cdot e^{-Q/2} \tag{3d. 1. 1}$$

のとき2変量正規分布に従うという．ただし $Q = \lambda_{11}(x-\xi)^2 + 2\lambda_{12}(x-\xi)(y-\eta) + \lambda_{22}(y-\eta)^2$ は正定符号である．

$$Q = \lambda_{22}\left(y - \eta + \frac{\lambda_{12}}{\lambda_{22}}\overline{x-\xi}\right)^2 + \left(\lambda_{11} - \frac{\lambda_{12}^2}{\lambda_{22}}\right)(x-\xi)^2$$

$$= \lambda_{22}u^2 + \left(\lambda_{11} - \frac{\lambda_{12}^2}{\lambda_{22}}\right)v^2 \tag{3d. 1. 2}$$

と書く．ただし u と v は (3d. 1. 2) で定義される．(3d. 1. 1) の Q に (3d. 1. 2) を代入し，u と v であらわすと u と v は独立な正規分布に従うことがわかる．これより，いくつかの結果が得られる．

1.　$v \sim N(0, \lambda_{22}/\Delta)$ である．ただし $\Delta = \lambda_{11}\lambda_{22} - \lambda_{12}^2$ である．$v = x - \xi$ だから $E(x) = \xi$, $V(x) = \lambda_{22}/\Delta = \sigma_1^2$（とおく）である．

2.　同様に対称性から $E(y) = \eta$, $V(y) = \lambda_{11}/\Delta = \sigma_2^2$ である．

3.　u と v の独立性から $0 = \mathrm{cov}(u, v) = \mathrm{cov}(x, y) + (\lambda_{12}/\lambda_{22})V(x) = 0$ となり，$\mathrm{cov}(x, y) = -\lambda_{12}/\Delta = \rho\sigma_1\sigma_2$ が得られる．ここで ρ は x と y の相関係数である．

4.　σ_1, σ_2, ρ であらわすと次のようになる．

$$Q = \frac{1}{1-\rho^2}\left[\frac{(x-\xi)^2}{\sigma_1^2} - \frac{2\rho(x-\xi)}{\sigma_1}\frac{(y-\eta)}{\sigma_2} + \frac{(y-\eta)^2}{\sigma_2^2}\right].$$

5.　x の周辺分布は，v の周辺分布と同じであるから $N(\xi, \sigma_1^2)$ である．同様に y の周辺分布は $N(\eta, \sigma_2^2)$ である．

6.　x を与えたときの y の条件付密度関数［(3d. 1. 1) を x の周辺密度関数で割る］は

$$c \cdot \exp\left\{\frac{-\lambda_{22}}{2}\left[y - \eta + \frac{\lambda_{21}}{\lambda_{22}}(x-\xi)\right]^2\right\}$$

$$= c \cdot \exp\left\{\frac{-1}{2\sigma_2^2(1-\rho^2)}\left[y - \eta - \frac{\rho\sigma_2}{\sigma_1}(x-\xi)\right]^2\right\}$$

である．これもまた正規分布で分散が $\sigma_2^2(1-\rho^2)$，平均（回帰関数）は

$$\eta + \frac{\rho\sigma_2}{\sigma_1}(x-\xi) = \alpha + \beta x, \qquad \beta = \frac{\rho\sigma_2}{\sigma_1}, \qquad \alpha = \eta - \beta\xi$$

で x の1次関数である．

7.　$\sigma_2^2(1-\rho^2) = \sigma_{2\cdot1}^2$ と書くと，(x, y) の密度関数は

$$(2\pi\sigma_1\sigma_{2\cdot1})^{-1}\exp\left[-\frac{(x-\xi)^2}{2\sigma_1^2} - \frac{(y-\alpha-\beta x)^2}{2\sigma_{2\cdot1}^2}\right] \tag{3d. 1. 3}$$

と書ける．

8.　(3d. 1. 3) から $x - \xi$ と $y - \alpha - \beta x$ は独立で，それらの同時特性関数は（1変量の場合の公式を用いると）

$$e^{-(t_1^2\sigma_1^2 + t_2^2\sigma_{2\cdot1}^2)/2}$$

となる．これから $x - \xi$ と $y - \eta$ の同時特性関数

$$e^{-(t_1^2\sigma_1^2 + 2\rho\sigma_1\sigma_2 t_1 t_2 + t_2\sigma_2^2)/2} \tag{3d. 1. 4}$$

が得られる．x と y の同時特性関数が必要ならば (3d. 1. 4) に $e^{it_1\xi + it_2\eta}$ をかければよい．

9.　(3d. 1. 4) で $ta = t_1$, $tb = t_2$ を代入すると $a(x-\xi) + b(y-\eta)$ の特性関数が求まり，平均が0，分散が $a^2\sigma_1^2 + 2\rho ab\sigma_1\sigma_2 + b^2\sigma_2^2$ の正規分布に従うことがわかる．

2変数 (x, y) は，x と y のあらゆる1次関数が正規分布に従うように分布しているとしよう．x と y の同時分布についてどのようなことがわかるだろうか？

一般性を失なうことなく $E(x) = E(y) = 0$, $V(x) = \sigma_1^2$, $V(y) = \sigma_2^2$, $\mathrm{cov}(x, y) = \rho\sigma_1\sigma_2$ とできる．分散が $\sigma^2 = a^2\sigma_1^2 + 2ab\rho\sigma_1\sigma_2 + b^2\sigma_2^2$ である1次関数 $ax + by$ を考えよう．その特性関数は仮定より

$$E[\exp it(ax + by)] = e^{-\sigma^2 t^2/2}$$

である．したがって $at=t_1$, $bt=t_2$ と書くと

$$E[\exp(it_1x+it_2y)]=e^{-(\sigma_1^2t_1^2+2\rho\sigma_1\sigma_2t_1t_2+\sigma_2^2t_2^2)/2} \tag{3d. 1. 5}$$

となる．(3d. 1. 5) はすべての a, b, したがってすべての t_1, t_2 にたいして成り立つ．一方 (3d. 1. 5) の右辺は (3d. 1. 4) のように2変量正規分布の特性関数である．(x, y) が2変量正規分布に従うならば，あらゆる1次関数も正規分布に従うことはすでに確かめたが，ここでその逆も証明された．

3d. 2 標本分布

$u_i\sim N(0,\sigma^2)$, $i=1,\cdots,n$ は n 個の独立な確率変数とし

$$\bar{u}=\frac{(u_1+\cdots+u_n)}{n}$$

$$S_{uv}=u_1(v_1-\bar{v})+\cdots+u_n(v_n-\bar{v})$$

$$S_{uu}=(u_1-\bar{u})^2+\cdots+(u_n-\bar{u})^2$$

$$S_{vv}=(v_1-\bar{v})^2+\cdots+(v_n-\bar{v})^2$$

を定義する．ただし $v_1,\cdots,v_n$ は定数である．$\bar{u}$ と S_{uv} が u_i の1次関数だから

$$\bar{u}\sim N\left(0,\frac{\sigma^2}{n}\right),\quad \text{したがって}\ \frac{n\bar{u}^2}{\sigma^2}\sim\chi^2(1) \tag{3d. 2. 1}$$

$$S_{uv}\sim N(0,\sigma^2S_{vv}),\quad \text{したがって}\ \frac{S_{uv}^2}{\sigma^2S_{vv}}\sim\chi^2(1) \tag{3d. 2. 2}$$

である．さらに $\mathrm{cov}(\bar{u},S_{uv})=\sum(v_i-\bar{v})\sigma^2/n=0$. よって $\bar{u}$ と S_{uv} は独立である．

$$S_{uu}=\sum u_i^2-n\bar{u}^2=\frac{S_{uv}^2}{S_{vv}}+\left(S_{uu}-\frac{S_{uv}^2}{S_{vv}}\right)$$

と書くと

$$\frac{\sum u_i^2}{\sigma^2}=\frac{n\bar{u}^2}{\sigma^2}+\frac{S_{uv}^2}{\sigma^2S_{vv}}+\frac{1}{\sigma^2}\left(S_{uu}-\frac{S_{uv}^2}{S_{vv}}\right) \tag{3d. 2. 3}$$

$$\chi^2(n)=\chi^2(1)+\chi^2(1)+\chi^2(n-2)$$

である．最後の項は（コーシー＝シュヴァルツの不等式あるいはその他の方法により）正であり，他の項の分布は (3d. 2. 1), (3d. 2. 2) よりわかっているから，最後の項の分布は [(iv), 3b. 4] を適用すれば導ける．(3d. 2. 3) の結果は他の多くの方法，たとえば $(u_1,\cdots,u_n)$ に直交変換を行う方法で証明できる．(3d. 2. 1)～(3d. 2. 3) により次の統計量はすべて独立である．

$$T_1=\frac{\sqrt{n}\,\bar{u}}{\sigma}\sim N(0,1),\qquad T_2=\frac{S_{uv}}{\sigma\sqrt{S_{vv}}}\sim N(0,1)$$

$$T_3=\frac{1}{\sigma^2}\left(S_{uu}-\frac{S_{uv}^2}{S_{vv}}\right)\sim\chi^2(n-2) \tag{3d. 2. 4}$$

次に $(v_1,\cdots,v_n)$ を $(u_1,\cdots,u_n)$ とは独立な確率変数と考えると $(v_1,\cdots,v_n)$ と T_1, T_2, T_3 とは独立である．$v_i\sim N(0,\sigma_0^2)$, $i=1,\cdots,n$ なる仮定の下で互いに独立な次の2つの統計量をつけ加えよう．

$$T_4=\frac{\sqrt{n}\,\bar{v}}{\sigma_0}\sim N(0,1),\qquad T_5=\frac{S_{vv}}{\sigma_0^2}\sim\chi^2(n-1). \tag{3d. 2. 5}$$

5つの統計量 T_1, T_2, T_3, T_4, T_5 はすべて独立である．次に2変量正規分布に従う標本からのいくつかの統計量の標本分布を導こう．

2変量正規分布に従う標本の平均値の同時分布　(x, y) が2変量正規分布からの観測値ならば，(3d. 1. 3) で

$$v=x-\xi\sim N(0,\sigma_1^2),$$

$$u=y-\eta-\beta(x-\xi)\sim N(0,\sigma_{2\cdot1}^2) \tag{3d. 2. 6}$$

であり，これらは独立であることが示されている．n 対の標本 (x, y) と n 対の (u, v) の関係は

$$\bar{v}=\bar{x}-\xi,\qquad \bar{u}=\bar{y}-\eta-\beta(\bar{x}-\xi)$$

である．したがって (3d. 2. 4) と (3d. 2. 5) から

$$T_1=\frac{[\bar{y}-\eta-\beta(\bar{x}-\xi)]\sqrt{n}}{\sigma_{2\cdot1}}$$

$$T_4=\frac{(\bar{x}-\xi)\sqrt{n}}{\sigma_1} \tag{3d. 2. 7}$$

は互いに独立で，それぞれ $N(0,1)$ に従うことがわかる．T_1 と T_4 の密度関数は $c\cdot e^{-(T_1^2+T_4^2)/2}$ である．変換 (3d. 2. 7) により T_1, T_4 を確率変数 $\bar{x}, \bar{y}$ に変換すると $\bar{x}$ と $\bar{y}$ の密度関数は $\mathrm{const}\cdot\exp(Q)$ となる．ここで

$$Q=-n\left\{\frac{(\bar{x}-\xi)^2}{2\sigma_1^2}+\frac{[\bar{y}-\eta-\beta(\bar{x}-\xi)]^2}{2\sigma_{2\cdot1}^2}\right\}$$

$$=\frac{-n}{2(1-\rho^2)}\left[\frac{(\bar{x}-\xi)^2}{\sigma_1^2}-\frac{2\rho(\bar{x}-\xi)(\bar{y}-\eta)}{\sigma_1\sigma_2}+\frac{(\bar{y}-\eta)^2}{\sigma_2^2}\right].$$

これは2変量正規分布である．

分散と共分散の同時分布　(u, v) と (x, y) の関係 (3d. 2. 6) を用いると

$$S_{vv}=S_{xx},\qquad S_{uv}=S_{xy}-\beta S_{xx}$$

$$S_{uu}-\frac{S_{uv}^2}{S_{vv}}=S_{yy}-\frac{S_{xy}^2}{S_{xx}}$$

であることがわかる．3つの統計量 T_2, T_3, T_5 を次のように用いる．

$$T_2=\frac{S_{uv}}{\sigma_{2\cdot1}\sqrt{S_{vv}}}=\frac{S_{xy}-\beta S_{xx}}{\sigma_{2\cdot1}\sqrt{S_{xx}}},$$

$$T_3=\frac{1}{\sigma_{2\cdot1}^2}\left(S_{uu}-\frac{S_{uv}^2}{S_{vv}}\right)=\frac{1}{\sigma_{2\cdot1}^2}\left(S_{yy}-\frac{S_{xy}^2}{S_{xx}}\right), \tag{3d. 2. 8}$$

$$T_5=\frac{S_{vv}}{\sigma_1^2}=\frac{S_{xx}}{\sigma_1^2}.$$

変換のヤコビアンは

$$\frac{D(T_2, T_3, T_5)}{D(S_{xx}, S_{xy}, S_{yy})} = \frac{1}{\sigma_1^2} \cdot \frac{1}{\sigma_{2\cdot1}\sqrt{S_{xx}}} \cdot \frac{1}{\sigma_{2\cdot1}^2}$$

である．T_2, T_3, T_5 の同時密度関数は ((3d. 2. 4)，(3d. 2. 5) から)

$$\mathrm{c}\cdot e^{-(T_2^2+T_3+T_5)/2} T_3^{(n-4)/2} T_5^{(n-3)/2} \tag{3d. 2. 9}$$

である．一方

$$T_2^2 + T_3 + T_5 = \frac{1}{1-\rho^2}\left(\frac{S_{xx}}{\sigma_1^2} - 2\rho\frac{S_{xy}}{\sigma_1\sigma_2} + \frac{S_{yy}}{\sigma_2^2}\right) = Q$$

である．よって(3d. 2. 9) を S_{xx}, S_{xy}, S_{yy} であらわし，ヤコビアンを導入すると標本分散と共分散の同時密度関数は

$$\mathrm{c}\cdot e^{-Q/2}(S_{xx}S_{yy} - S_{xy}^2)^{(n-4)/2} \tag{3d. 2. 10}$$

となる．

相関係数の分布　(3d. 2. 10) の分布を用いると Fisher(1915) が彼の最初の論文で示したように2つの適当に選んだ関数に関して積分することにより $r = S_{xy} \div \sqrt{S_{xx}S_{yy}}$ の分布を求めることができる．しかし，また，基本的な統計量 T_2, T_3, T_5 を用いて直接 r の分布をきめることもできる．

r を導入し，$\sigma_{2\cdot1}^2 = \sigma_2^2(1-\rho^2)$ と書くと (3d. 2. 8) の関係は

$$T_2 = \frac{r\sqrt{S_{yy}}}{\sigma_2\sqrt{1-\rho^2}} - \frac{\rho\sqrt{S_{xx}}}{\sigma_1\sqrt{1-\rho^2}},$$

$$T_3 = \frac{S_{yy}(1-r^2)}{\sigma_2^2(1-\rho^2)}, \qquad T_5 = \frac{S_{xx}}{\sigma_1^2}$$

となる．T_3 と T_5 を用いて T_2 から S_{xx} と S_{yy} を消去するとフレイザー＝スプロットの基本公式

$$T_2 = \frac{r}{\sqrt{1-r^2}}\sqrt{T_3} - \frac{\rho}{\sqrt{1-\rho^2}}\sqrt{T_5} \tag{3d. 2. 11}$$

が得られる．(3d. 2. 11) を用いると T_2, T_3, T_5 の同時分布から r の分布が得られる．$\rho = 0$ のとき

$$T_2 = \frac{r}{\sqrt{1-r^2}}\sqrt{T_3}$$

となり $r/\sqrt{1-r^2}$ の分布は $T_2/\sqrt{T_3}$ の分布に等しい．一方

$$T_2/\sqrt{T_3 \div (n-2)} \sim S(n-2)$$

すなわちスチューデントの分布 (3b. 1. 8) に従う．これより

$$\frac{r\sqrt{n-2}}{\sqrt{1-r^2}} \sim S(n-2) \tag{3d. 2. 12}$$

となり，r の有意性は t 分布のパーセント点の表を用いて検定できる．(3d. 2. 12) から変数変換により r の分布は

$$\frac{\Gamma[(n-1)/2]}{\sqrt{\pi}\,\Gamma[(n-2)/2]}(1-r^2)^{(n-4)/2}dr, \qquad -1 \leqslant r \leqslant 1 \tag{3d. 2. 13}$$

であることがわかる．

$\rho \neq 0$ のときの分布をきめるには T_2, T_3, T_5 の同時分布を用いる．$t = r/\sqrt{1-r^2}$，$\theta = \rho/\sqrt{1-\rho^2}$ とおくと (3d. 2. 11) は

$$t\sqrt{T_3} = T_2 + \theta\sqrt{T_5} \quad \text{すなわち} \quad t = \frac{T_2 + \theta\sqrt{T_5}}{\sqrt{T_3}} \tag{3d. 2. 14}$$

となる．T_2, T_3, T_5 の同時密度関数は

$$\mathrm{c}\cdot e^{-(T_2^2+T_3+T_5)/2} T_3^{(n-4)/2} T_5^{(n-3)/2} \tag{3d. 2. 15}$$

だから，問題は (3d. 2. 14) で定義した T_2, T_3, T_5 の関数である t の分布を求めることに帰着する．

$$T_2 = t\sqrt{T_3} - \theta\sqrt{T_5}, \qquad dT_2 = \sqrt{T_3}dt$$

を (3d. 2. 15) に代入すると t, T_3, T_5 の同時密度関数は

$$\mathrm{c}\cdot e^{-[T_3(1+t^2)+T_5(1+\theta^2)-2t\theta\sqrt{T_3T_5}]/2}(T_3T_5)^{(n-3)/2}$$

となる．T_3, T_5 を $u = T_3(1+t^2)$，$v = T_5(1+\theta^2)$ に変換すると t, u, v の同時密度関数

$$\mathrm{c}\cdot\left(\frac{1}{1+t^2}\right)^{(n-1)/2} e^{-(u+v-2\rho r\sqrt{uv})/2}(uv)^{(n-3)/2}$$

と r, u, v $(t = r/\sqrt{1-r^2})$ の同時密度関数

$$\mathrm{c}\cdot(1-r^2)^{(n-4)/2}e^{-(u+v-2\rho r\sqrt{uv})/2}(uv)^{(n-3)/2}$$

が得られる．r のみの密度関数は，積分

$$\mathrm{c}\cdot(1-r^2)^{(n-4)/2}\int_0^\infty\int_0^\infty e^{-(u+v-2\rho r\sqrt{uv})/2}(uv)^{(n-3)/2}dudv$$

であらわすことができる．

$e^{\rho r\sqrt{uv}}$ を展開し，項別積分すると，$\rho \neq 0$ のときの r の分布は

$$\mathrm{c}\cdot(1-r^2)^{(n-4)/2}\sum_{s=0}^{\infty}\Gamma^2\left(\frac{n+s-1}{2}\right)\frac{(2\rho r)^s}{s!} \tag{3d. 2. 16}$$

となる．ただし定数 c は

$$\frac{2^{n-3}}{\pi(n-3)!}(1-\rho^2)^{(n-1)/2}$$

である．

回帰係数の分布　y の x への回帰係数は $b = S_{xy}/S_{xx}$ で推定する．b の分布を求めるのに関係のある統計量は

$$T_2 = \frac{b-\beta}{\sigma_{2\cdot1}}\sqrt{S_{xx}} = \frac{\sigma_1}{\sigma_{2\cdot1}}\sqrt{T_5}(b-\beta),$$

$$T_3 = \frac{1}{\sigma_{2\cdot1}^2}\left(S_{yy} - \frac{S_{xy}^2}{S_{xx}}\right), \qquad T_5 = \frac{S_{xx}}{\sigma_1^2} \tag{3d. 2. 17}$$

である．最も役に立つ分布は，統計量 t の分布

$$t = \frac{(b-\beta)\sqrt{S_{xx}}}{\left[\left(S_{yy} - \frac{S_{yx}^2}{S_{xx}}\right)\Big/(n-2)\right]^{1/2}} = \frac{T_2}{\sqrt{T_3 \div (n-2)}} \sim S(n-2)$$

つまり自由度 $n-2$ のスチューデントの分布である．推定値の仮説の値 β からのズレの観測値の有意性は t 統計量で検定する．しかし b のみの分布を求めるならば，同時密度関数が

$$\mathrm{c}\cdot e^{-(T_2^2+T_3)/2}T_5^{(n-3)/2}$$

である統計量 T_2 と T_5 を考えなければならない．(3d. 2. 17) の関係により b と T_5 に変換すると密度関数は

$$\mathrm{c}\cdot e^{-T_5[\sigma_{2\cdot1}^2+\sigma_1^2(b-\beta)^2]/2}T_5^{(n-2)/2}$$

となる．T_5 について積分すれば，b のみの密度関数は

$$\mathrm{c}\cdot[\sigma_{2\cdot1}^2 + \sigma_1^2(b-\beta)^2]^{-n/2} \tag{3d. 2. 18}$$

となる．

補足と問題

1　1次形式と2次形式の独立性　$y_i \sim N(\mu_i, \sigma^2)$，$i = 1, \cdots, n$ はすべて独立とし，$\mathbf{Y}' = (y_1, \cdots, y_n)$，$\boldsymbol{\mu}' = (\mu_1, \cdots, \mu_n)$ と書く．以下の結果を証明せよ．

1.1　$\mathbf{Y'AY}$ と $\mathbf{Y'BY}$ が互いに独立な分布に従うための十分条件は $\mathbf{AB} = \mathbf{BA} = 0$ である．［ヒント：$\mathbf{A}$ と $\mathbf{B}$ のスペクトル分解 ((1c. 3. 4), (1c. 3. 5)) により

$$\mathbf{A} = \lambda_1\mathbf{P}_1\mathbf{P}_1' + \cdots + \lambda_r\mathbf{P}_r\mathbf{P}_r', \qquad \lambda_i \neq 0,\ i = 1, \cdots, r$$
$$\mathbf{B} = \nu_1\mathbf{Q}_1\mathbf{Q}_1' + \cdots + \nu_s\mathbf{Q}_s\mathbf{Q}_s', \qquad \nu_j \neq 0,\ j = 1, \cdots, s$$

となる．ただし $r = \mathrm{rank}\,\mathbf{A}$，$s = \mathrm{rank}\,\mathbf{B}$ で $\mathbf{P}_i$ は互いに正規直交し，$\mathbf{Q}_j$ も互いに正規直交する．$\mathbf{AB} = 0 = \sum\sum \lambda_i\nu_j\mathbf{P}_i\mathbf{P}_i'\mathbf{Q}_j\mathbf{Q}_j'$ である．$\mathbf{P}_r'$ を前から，$\mathbf{Q}_s$ を後からかけると $\mathbf{P}_r'\mathbf{Q}_s = 0$ となる．したがって $\mathbf{P}_i$ と $\mathbf{Q}_j$ はすべて正規直交する．すると $\mathbf{P}_i'\mathbf{Y}$ と $\mathbf{Q}_j'\mathbf{Y}$ はすべて独立で正規分布に従う．条件の十分性は $\mathbf{Y'AY} = \lambda_1(\mathbf{P}_1'\mathbf{Y})^2 + \cdots + \lambda_r(\mathbf{P}_r'\mathbf{Y})^2$ と $\mathbf{Y'BY} = \nu_1(\mathbf{Q}_1'\mathbf{Y})^2 + \cdots + \nu_s(\mathbf{Q}_s'\mathbf{Y})^2$ が独立な確率変数の排反な集合に依存していることから証明できる．$\mathbf{Y'AY}$ と $\mathbf{Y'BY}$ はカイ2乗分布に従う必要はないことに注意せよ．］

1.2　$\mathbf{BY}$ を m 個の $\mathbf{Y}$ の1次関数とする．つまり $\mathbf{B}$ は $m \times n$ 行列である．$\mathbf{BY}$ が2次形式 $\mathbf{Y'AY}$ と独立な分布に従うための十分条件は $\mathbf{BA} = 0$ である．

［問題 **1.1** のように $\mathbf{A}$ のスペクトル分解を用いよ．条件 $\mathbf{BA} = 0$ は $\mathbf{BP}_i = 0$，$i = 1, \cdots, r$ を意味している．これより $\mathbf{BY}$ と1次関数 $\mathbf{P}_1'\mathbf{Y}, \cdots, \mathbf{P}_r'\mathbf{Y}$ は独立である．一方 $\mathbf{Y'AY} = \sum \lambda_i(\mathbf{P}_i'\mathbf{Y})^2$ である．］

1.3　条件 $\mathbf{BA} = 0$ はまた問題 **1.1** と問題 **1.2** の結果が成り立つための必要条件である．

［必要性の証明はやや複雑で特性関数を用いる．Ogawa(1949) を参照せよ．3b. 4 で2次形式の行列が対称でベキ等であるという条件を付加したとき，必要性が証明されている．］

1.4　$\mathbf{Y'AY}$ の特性関数は $|\mathbf{I} - 2i\sigma^2 t\mathbf{A}|^{-1/2}\exp[it\boldsymbol{\mu}'(\mathbf{I} - 2i\sigma^2 t\mathbf{A})^{-1}\mathbf{A}\boldsymbol{\mu}]$ で，$\mathbf{Y'AY}$ と $\mathbf{Y'BY}$ の同時特性関数は $|\mathbf{I} - 2i\sigma^2 t_1\mathbf{A} - 2i\sigma^2 t_2\mathbf{B}|^{-1/2}\exp[\boldsymbol{\mu}'(\mathbf{I} - 2i\sigma^2 t_1\mathbf{A} - 2i\sigma^2 t_2\mathbf{B})^{-1}(it_1\mathbf{A} + it_2\mathbf{B})\boldsymbol{\mu}]$ であることを示せ．

1.5　**坂元－クレーグの定理**　$\mathbf{Y'AY}$ と $\mathbf{Y'BY}$ が独立な分布に従うならばすべての t_1, t_2 にたいして $|\mathbf{I} - 2i\sigma^2 t_1\mathbf{A} - 2i\sigma^2 t_2\mathbf{B}| = |\mathbf{I} - 2i\sigma^2 t_1\mathbf{A}|\,|\mathbf{I} - 2i\sigma^2 t_2\mathbf{B}|$ である．これから $\mathbf{AB} = 0$ を導け．

2　n 個の独立な確率変数についての 3b. 5 の模型

$$y_i \sim N(x_{i1}\beta_1 + \cdots + x_{im}\beta_m, \sigma^2), \qquad i = 1, \cdots, n$$

を考える．$\mathbf{Y}' = (y_1, \cdots, y_n)$，$\boldsymbol{\beta}' = (\beta_1, \cdots, \beta_m)$，$\mathbf{X} = (x_{ij})$（階数は r），$(\mathbf{X'X})^-$ を $\mathbf{X'X}$ の一般逆行列 (1b. 5) とする．さらに $\hat{\boldsymbol{\beta}}$ が方程式 $\mathbf{X'X}\boldsymbol{\beta} = \mathbf{X'Y}$ の解となるように $\hat{\boldsymbol{\beta}} = (\mathbf{X'X})^-\mathbf{X'Y}$ とする．

2.1　次の (a)，(b)，(c) を証明せよ．

(a)　$\mathbf{X} = \mathbf{X}(\mathbf{X'X})^-\mathbf{X'X}$，(b)　行列 $\mathbf{I} - \mathbf{X}(\mathbf{X'X})^-\mathbf{X}'$ はベキ等である．［ヒント：$\mathbf{G} = \mathbf{X}[\mathbf{I} - (\mathbf{X'X})^-\mathbf{X'X}]$ とおく．$\mathbf{G'G}$ を計算しそれが 0 であることを示せ．これより $\mathbf{G} = 0$ が得られる．(b) を証明するために $\mathbf{I} - \mathbf{X}(\mathbf{X'X})^-\mathbf{X}'$ の2乗を計算し (a) の結果を用いよ．］(c)　さらに $\mathbf{X}' = \mathbf{X'X}(\mathbf{X'X})^-\mathbf{X}'$．

2.2　$\mathbf{H}$ を $\mathscr{M}(\mathbf{H}) = \mathscr{M}(\mathbf{X}')$ となるような $m \times k$ 行列とする．$\mathbf{Z} = \mathbf{H}'\hat{\boldsymbol{\beta}}$ と $R_0^2 = (\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})'(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})$ は互いに独立な分布に従うことを示せ．［ヒント：$\mathscr{M}(\mathbf{H}) = \mathscr{M}(\mathbf{X}')$ だから $\mathbf{H} = \mathbf{X}'\mathbf{C}$ となるような行列 $\mathbf{C}$ が存在する．したがって

$$\mathbf{H}'\hat{\boldsymbol{\beta}} = \mathbf{C}'\mathbf{X}(\mathbf{X'X})^-\mathbf{X'Y} = \mathbf{BY},$$
$$R_0^2 = \mathbf{Y'Y} - \mathbf{Y'X}(\mathbf{X'X})^-\mathbf{X'Y} = \mathbf{Y}'[\mathbf{I} - \mathbf{X}(\mathbf{X'X})^-\mathbf{X}']\mathbf{Y} = \mathbf{Y'AY}$$

である．われわれは $\mathbf{Y}$ の1次関数の集合と2次関数が独立であることを証明したい．$\mathbf{B}$ と $\mathbf{A}$ の2つの項の積は問題 **2.1** の (c) により 0 であるから

$$\mathbf{BA} = [\mathbf{C}'\mathbf{X}(\mathbf{X'X})^-\mathbf{X}'][\mathbf{I} - \mathbf{X}(\mathbf{X'X})^-\mathbf{X}'] = 0$$

である．問題 **1.2** の十分条件から求める結果が得られる．］

2.3　$\mathbf{Z} = \mathbf{H}'\hat{\boldsymbol{\beta}}$ は平均が $\mathbf{H}'\boldsymbol{\beta}$，分散共分散行列が $\sigma^2\mathbf{D}$ の k 変量正規分布に従う．ただし $\mathbf{D} = \mathbf{H}'(\mathbf{X'X})^-\mathbf{H}$．また $R_0^2 \sim \sigma^2\chi^2(n-r)$ である．［ヒント：$\mathbf{Z} = \mathbf{C}'\mathbf{X}(\mathbf{X'X})^-\mathbf{X'Y} = \mathbf{BY}$，すなわち正規確率変数の1次関数である．(3b. 3. 4) を適用すると $\mathbf{Z}$ は k 変量正規分布に従う．

$$R_0^2 = \mathbf{Y}'[\mathbf{I} - \mathbf{X}(\mathbf{X'X})^-\mathbf{X}']\mathbf{Y}$$

を考えよう．行列 $\mathbf{I} - \mathbf{X}(\mathbf{X'X})^-\mathbf{X}'$ はベキ等である．したがって $R_0^2 \sim \sigma^2\chi^2(n-r)$ である．ただし $n - r = \mathrm{trace}\,\mathbf{I} - \mathrm{trace}\,\mathbf{X}(\mathbf{X'X})^-\mathbf{X}' = n - \mathrm{trace}(\mathbf{X'X})^-\mathbf{X'X} = n - \mathrm{rank}\,\mathbf{X'X} = n - \mathrm{rank}\,\mathbf{X}$ である．さらに問題 **1.2** の十分条件を適用すれば $\mathbf{Z}$ と R_0^2 は互いに独立な分布に従う．］

2.4　$\boldsymbol{\beta}^*, \boldsymbol{\lambda}^*$ を次の方程式の解とする．

$$\mathbf{X'X}\boldsymbol{\beta} + \mathbf{H}\boldsymbol{\lambda} = \mathbf{X'Y}$$
$$\mathbf{H}'\boldsymbol{\beta} = \boldsymbol{\theta}_0 \qquad \text{(所与の値)}$$

この方程式は条件 $\mathbf{H}'\boldsymbol{\beta} = \boldsymbol{\theta}_0$ の下で $(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})$ を $\boldsymbol{\beta}$ に関して微分することによって得られる．

$$R_1^2 = (\mathbf{Y} - \mathbf{X}\beta^*)'(\mathbf{Y} - \mathbf{X}\beta^*)$$
$$= (\mathbf{Z} - \theta_0)'\mathbf{D}^{-1}(\mathbf{Z} - \theta_0) + R_0^2$$

であることを示せ．ただし $\mathbf{Z}$ と $\mathbf{D}$ は問題 2.3 で定義されている．［ヒント：$R_1^2 = (\mathbf{Y} - \mathbf{X}\hat{\beta})'(\mathbf{Y} - \mathbf{X}\hat{\beta}) + (\hat{\beta} - \beta^*)'\mathbf{X}'\mathbf{X}(\hat{\beta} - \beta^*)$ だから第2項が $(\mathbf{Z} - \theta_0)'\mathbf{D}^{-1}(\mathbf{Z} - \theta_0)$ であることを示せば十分である．$\mathscr{M}(\mathbf{H}) \subset \mathscr{M}(\mathbf{X}'\mathbf{X})$ だから $\mathbf{H} = \mathbf{X}'\mathbf{X}\mathbf{C}$ である．β^*, λ^* にたいする方程式から $\mathbf{X}'\mathbf{X}(\hat{\beta} - \beta^*) = \mathbf{H}\lambda^* = \mathbf{X}'\mathbf{X}\mathbf{C}\lambda^*$，したがって $(\hat{\beta} - \beta^*)'\mathbf{X}'\mathbf{X}(\hat{\beta} - \beta^*) = \lambda^{*\prime}\mathbf{C}'\mathbf{X}'\mathbf{X}\mathbf{C}\lambda^*$ となる．また方程式から $\mathbf{C}'\mathbf{X}'\mathbf{X}\beta^* - \mathbf{H}'\beta^* + \mathbf{C}'\mathbf{H}\lambda^* = \mathbf{C}'\mathbf{X}'\mathbf{Y} - \theta_0$ すなわち $\mathbf{C}'\mathbf{H}\lambda^* = \mathbf{Z} - \theta_0$，つまり $\mathbf{C}'\mathbf{X}'\mathbf{X}\mathbf{C}\lambda^* = \mathbf{Z} - \theta_0 \Rightarrow \lambda^* = \mathbf{D}^{-1}(\mathbf{Z} - \theta_0)$ となる．ここで問題 2.3 の $\mathbf{D}$ にたいする式の $\mathbf{H}$ に〔$\mathbf{H} = \mathbf{X}'\mathbf{X}\mathbf{C}$ を〕代入すると $\mathbf{D} = \mathbf{C}'\mathbf{X}'\mathbf{X}\mathbf{C}$ となる．これより結果が得られる．］

［注意　問題 2.1 から 2.4 までで β に制約のないときあるいは $\mathbf{H}'\beta = \theta_0$ という制約があるとき，それぞれ R_0^2 あるいは R_1^2 がそのときの最小値に対応しているならば，最小2乗法の最初の2つの基本定理の別の証明法が得られたことになる．］

2.5　3b.5 の (ii) で条件 $\mathscr{M}(\mathbf{H}) \subset \mathscr{M}(\mathbf{X}')$ をゆるめると $R_1^2 - R_0^2$ は自由度が t の非心カイ2乗分布に従う．ただし $t = \mathrm{rank}[\mathscr{M}(\mathbf{H}) \cap \mathscr{M}(\mathbf{X}')]$ である．非心率を求めよ．

2.6　R_0^2, R_1^2, R_2^2 をそれぞれ β に制約のないときの，$\mathbf{H}'\beta = \xi$ という制約があるときの，および $\mathbf{H}'\beta = \xi$，$\mathbf{G}'\beta = \eta$ という制約があるときの $(\mathbf{Y} - \mathbf{X}\beta)'(\mathbf{Y} - \mathbf{X}\beta)$ の最小値とする．このとき R_0^2，$R_1^2 - R_0^2$，$R_2^2 - R_1^2$ は互いに独立な分布に従う．$R_0^2 \sim \sigma^2\chi^2(n - r)$ で $R_1^2 - R_0^2$ と $R_2^2 - R_1^2$ はそれぞれ自由度 k，自由度 t の非心カイ2乗分布に従う．ただし $r = \mathrm{rank}\,\mathbf{X}$，$k = \mathrm{rank}[\mathscr{M}(\mathbf{H}) \cap \mathscr{M}(\mathbf{X}')]$，$t + k = \mathrm{rank}[\mathscr{M}(\mathbf{H}) + \mathscr{M}(\mathbf{G}) \cap \mathscr{M}(\mathbf{X}')]$ である．制約が成り立つならば非心率は0である．

2.7　問題 2.6 の結果を，適当に定義した $R_0^2, R_1^2, R_2^2, R_3^2, \cdots$ に一般化せよ．

2.8　共通の分散が σ^2，平均ベクトルが μ である n 個の独立な正規確率変数のベクトル $\mathbf{Y}$ を考えよう．$\omega_1, \omega_2, \cdots, \omega_k$ を n 次元ユークリッド空間の部分空間とする．$R_1^2, R_2^2, \cdots, R_k^2$ をそれぞれ $\mu \in \omega_1$，$\mu \in \omega_1 \cap \omega_2$，$\cdots$，$\mu \in \omega_1 \cap \omega_2 \cap \cdots \cap \omega_k$ という制約があるときの $(\mathbf{Y} - \mu)'(\mathbf{Y} - \mu)$ の最小値とする．R_1^2，$R_2^2 - R_1^2$，$R_3^2 - R_2^2$，$\cdots$ は自由度が $k_1 = n - d[\omega_1]$，$k_2 = d[\omega_1] - d[\omega_1 \cap \omega_2]$，$k_3 = d[\omega_1 \cap \omega_2] - d[\omega_1 \cap \omega_2 \cap \omega_3]$，$\cdots$ の互いに独立な非心カイ2乗確率変数である．ただし d は指定した部分空間の次元をあらわす．この問題は問題 2.6 の特別な場合であり，制約が同次である．［ヒント：$\omega_1 \cap \cdots \cap \omega_k$ の正規直交基底を選ぶことにより，直交変換を構成し，ベクトルを加えることにより $\omega_1 \cap \cdots \cap \omega_{k-1}$ の正規直交基底をつくれ．これを全体の空間の正規直交基底が得られるまで行え．］

2.9　問題 2.8 で $R_i^2 - R_{i-1}^2$ が $\mu \in \omega_i$ という制約の下で $(\mathbf{Y} - \mu)'(\mathbf{Y} - \mu)$ の最小値になるための条件を述べよ．

2.10　列ベクトル β が $\mathbf{H}'\beta = 0$（$\mathbf{H}'$ は $k \times m$ 行列）という条件の制約をうけるときベクトル $\mathbf{X}\beta$（$\mathbf{X}$ は $n \times m$ 行列）によって生成される各部分空間の次元は $d\{\mathscr{M}(\mathbf{X}')\} - d\{\mathscr{M}(\mathbf{X}') \cap \mathscr{M}(\mathbf{H})\}$ であることを示せ．

2.11　$\mathbf{Y}$ が，平均 $\mathbf{X}\beta$，分散共分散行列 Λ の n 変量正規分布に従い，$R_0^2, R_1^2, \cdots$ は $(\mathbf{Y} - \mathbf{X}\beta)'\Lambda^{-1}(\mathbf{Y} - \mathbf{X}\beta)$ の最小値であると定義するならば，3b.5 の結果と問題 2.5 から 2.9 までの結果は同じように成り立つことを確かめよ．

3　$x_i \sim N(\mu, \sigma^2)$，$i = 1, \cdots, n$ を n 個の独立な確率変数とする．$m_1, \cdots, m_n$ が一定の値であるならば，次の相関係数の分布を求めよ．

$$r = \sum x_i(m_i - \overline{m}) \div [\sum(m_i - \overline{m})^2 \cdot \sum(x_i - \bar{x})^2]^{1/2}$$

［ヒント：適当な変数変換あるいはその他の方法により $r = x/\sqrt{x^2 + y}$ であることを示せ．ただし $x \sim N(0, 1)$ と $y \sim \chi^2(n - 2)$ は互いに独立である．次に，(3b.1.13) を用いよ．］

4　ラプラス分布　ラプラス分布の密度関数は

$$L(x|\mu, \lambda) = (2\lambda)^{-1}\exp(-|x - \mu|/\lambda)$$

という形である．特性関数は $(1 + \lambda^2t^2)^{-1}\exp(it\mu)$ であることを示せ．この分布はあらゆる次数の積率が存在し，たたみこみに関して閉じていない．

5　パレート分布　ある確率変数の密度関数が

$$\frac{\alpha}{x_0}\left(\frac{x_0}{x}\right)^{\alpha+1}, \qquad x > x_0,\ \alpha > 0$$

でそれ以外では0である．この分布の積率を調べよ．平均は $\alpha > 1$ のときにのみ有限であり，$\alpha x_0/(\alpha - 1)$ であることを示せ．分布関数は $1 - (x_0/x)^\alpha$ という簡単な形である．これよりメディアンは $x_0 2^{1/\alpha}$ であることを示せ．ある限界 x_0 をこえる個人の所得の分布（打ち切られた分布）はしばしばパレートの形をとる．

6　対数正規分布　$\xi_1, \xi_2, \cdots$ の順に一連の独立でランダムな衝撃がおこり，ξ_r は生物の瞬間的な大きさを

$$x_{r+1} = x_r + \xi_r x_r \Rightarrow \frac{x_{r+1}}{x_r} = 1 + \xi_r = \eta_r$$

という関係で x_r から x_{r+1} に増大させる効果をもつとしよう．$w = \log x - \log x_0$——最後の段階と最初の段階の大きさの対数の差——ならば

$$w = \log \eta_0 + \cdots + \log \eta_{n-1}$$

となる．ただし $\log \eta_i$ は互いに独立である．各衝撃は微小な成長のみを与え，n は大きいものとする．［(iii)，3a.1］のように $V(w)$ は有限のままで $n \to \infty$ とする．極限では $w = \log x - \log x_0$ は正規分布に従い，x は分布

$$\frac{1}{x\sigma\sqrt{2\pi}}e^{-(\log x - \mu)^2/2\sigma^2}$$

に従う．x を α から測定したとすると，より一般的な形

$$\frac{1}{\sigma(x - \alpha)\sqrt{2\pi}}e^{-[\log(x-\alpha) - \mu]^2/2\sigma^2}, \qquad x > \alpha$$

が得られる．対数正規分布はしばしば所得の分布によくあてはまる．対数正規分布の平均と分散を求めよ．

7　ロジスティック分布　分布関数は

$$F(x) = \frac{1}{1 + e^{-(\alpha+\beta x)}}, \qquad \beta > 0$$

で与えられる．x の特性関数とモーメントを求めよ．これは生物検定法の問題で，耐性の〔個体変異の〕分布として用いられる．

8　ピアソン系の度数分布　$y = f(x)$ が密度関数をあらわすとする．カール・ピアソンは一般的な考察によって次の微分方程式を得た．

$$\frac{1}{y}\frac{dy}{dx}=\frac{x+a}{b_0+b_1x+b_2x^2}.$$

第3章で考えた分布のほとんどはこの系の特別な場合であり，定数 a, b_0, b_1, b_2 を特別に選ぶことにより導くことができる．分布の4次までのモーメントが存在するならば，定数はそれらであらわすことができることを示せ．

9　寿命試験のモデル　$F(t)$ をある構造が時刻 t までにこわれる確率とし，$\lambda(t)\Delta t+o(\Delta t)$ を t まで残存していたものが微小な時間間隔 Δt でこわれる確率とする．

9.1　$F(t)$ が関数方程式

$$F(t)=1-\exp\left[-\int\lambda(t)dt\right]$$

となるような微分方程式

$$\frac{f(t)}{1-F(t)}=\lambda(t)$$

をみたすことを証明せよ．ただし $F'(t)=f(t)$ は密度関数とする．

9.2　$\lambda(t)=c$（定数）という特別な場合について，指数分布

$$F(t)=1-e^{-ct}\quad すなわち\quad f(t)=ce^{-ct}$$

を導け．

9.3　指数分布は，時刻 T まで構造が残存していたとすると残りの寿命の条件付分布は T にはよらず初めからの寿命の分布と同一であるという性質によって特徴づけられることを示せ．［電気のヒューズや時計の石をはめた軸受のようにこの性質をもつ構造はいくつかある．］

9.4　$\lambda(t)=c\alpha t^{\alpha-1}$，$\alpha>1$，$c>0$ という特別な場合，この分布はワイブル分布とよばれる．分布関数は $1-\exp(-ct^{\alpha})$ であることを示せ．

9.5　マケハムの法則と呼ばれる $\lambda(t)=\gamma+\delta e^{\mu t}$ のときの $F(t)$ の式を導け．このようにして導かれた関数 $F(t)$ は死亡表をつくるのに使われる．

10　順序統計量　$x_1,\cdots,x_n$ を連続な分布関数 $F(x)$ からの n 個の独立な観測値とし，$y_1,\cdots,y_n$ を順序がついた観測値（いわゆる順序統計量）とする．x の密度関数が存在するならば $f(x)$ と書く．

10.1　y_1 の分布関数は $1-[1-F(y_1)]^n$ である．［ヒント：最小値が y_1 より小さい確率はすべての観測値が y_1 より小さくない確率を1からひいたものに等しい．］　$f(x)$ が存在するならば y_1 の密度関数は

$$n[1-F(y_1)]^{n-1}f(y_1)$$

である．

10.2　同様に最大値 y_n の分布関数と密度関数は

$$[F(y_n)]^n,\qquad n[F(y_n)]^{n-1}f(y_n)$$

である．

10.3　y_1 と y_n の同時分布関数は

$$[F(y_n)]^n-[F(y_n)-F(y_1)]^n$$

である．第1項は最大値が y_n より小さい確率をあらわし，第2項はすべての観測値が y_1 と y_n の間にある確率をあらわしていることに注意せよ．密度関数は，分布関数を微分すると

$$n(n-1)[F(y_n)-F(y_1)]^{n-2}f(y_1)f(y_n)$$

となる．これは範囲 y_n-y_1 の分布を求める出発点となる．

10.4　y_k よりも小さい観測値の数が k または $k+1$ または $k+2,\cdots$，または n ならば k 番目の順序統計量が y_k よりも小さい．その確率は

$$1-\sum_{r=0}^{k-1}\binom{n}{r}[F(y_k)]^r[1-F(y_k)]^{n-r}\qquad（2項定理により）$$

$$=1-\frac{n!}{(k-1)!(n-k)!}\int_0^{1-F}x^{n-k}(1-x)^{k-1}dx$$

である．これは重要な結果である．密度関数は

$$\frac{n!}{(k-1)!(n-k)!}[1-F(y_k)]^{n-k}[F(y_k)]^{k-1}f(y_k)$$

である．メディアンやその他の役に立つ分位の分布はこの特別な場合として得られる．

10.5　$F(y_k)$ の分布は $B(k, n-k+1)$ であることを示せ．

10.6　順序統計量 $y_k, y_s, \cdots, y_u$ $(k<s<\cdots<u)$ の同時密度関数は

$$\frac{\Gamma(n+1)}{\Gamma(k)\Gamma(s-k)\cdots\Gamma(n-u+1)}[F(y_k)]^{k-1}[F(y_s)-F(y_k)]^{s-k-1}$$
$$\cdots[1-F(y_u)]^{n-u}f(y_k)\cdots f(y_u)$$

であることを示せ．

11　ディリクレ分布　$x_1,\cdots,x_{k+1}$ を母数 $(\alpha, p_1),\cdots,(\alpha, p_{k+1})$ の独立なガンマ確率変数とする．

$$y_i=x_i/(x_1+\cdots+x_i),\qquad i=2,\cdots,k+1$$

に独立な分布に従うことを証明せよ．さらに

$$z_i=x_i/(x_1+\cdots+x_{k+1}),\qquad i=1,\cdots,k$$

の同時密度関数は

$$\frac{\Gamma(p_1+\cdots+p_{k+1})}{\Gamma(p_1)\cdots\Gamma(p_{k+1})}z_1^{p_1-1}\cdots z_k^{p_k-1}(1-z_1-\cdots-z_k)^{p_{k+1}-1}$$

であり，ディリクレ分布と呼ばれる．

12　x_1, x_2 が $N(0,1)$ からの2つの独立な観測値とすると $z=x_1/x_2$ の分布を求めよ．

13　$x=x_1+\cdots+x_n$ を長方形分布 $R(0,1)$ からの n 個の独立な観測値の和とする．

$$P(x<X)=\frac{X^n}{n!0!}-\frac{(X-1)^n}{(n-1)!1!}+\frac{(X-2)^n}{(n-2)!2!}\cdots+(-1)^r\frac{(X-r)^n}{(n-r)!r!},\qquad r=[X]$$

を証明せよ（Rao(1942a)）．

14　$f(x)$ を確率変数 x の密度関数とすると $\alpha f(\alpha y)$ は $y=x/\alpha$ の密度関数である．密度関数が $\alpha_1 f(\alpha_1 y_1)$，$\alpha_2 f(\alpha_2 y_2)$ である確率変数 y_1, y_2 を考える．$y=y_1+y_2$ の密度関数がある α に関して同一の形 $\alpha f(\alpha y)$ となるとき，関数形 f を求めよ．f の2次までのモーメントが存在するならば，f は正規分布であることを導け［Polya (1923)］．

15　長さが1の円周上で，各長さが x である n 個の弧がランダムに区分される．（a）円周上のあらゆる点が少なくとも1つの弧に含まれる確率は

$$1-\binom{n}{1}(1-x)^{n-1}+\binom{n}{2}(1-2x)^{n-1}-\cdots\pm\binom{n}{k}(1-kx)^{n-1},\qquad k=\left[\frac{1}{x}\right]$$

であることを示せ．（b）間隙が i 個ある確率は

$$\binom{n}{i}\Big[(1-ix)^{n-1}-(n-i)[1-(i+1)x]^{n-1}+\cdots \pm\frac{(n-i)!}{(k-i)!(n-k)!}(1-kx)^{n-1}\Big]$$

である [Stevens (1939)].

16　x_r, $r=1,\cdots,n$ が自由度2のカイ2乗分布に従い，x_r は互いに独立ならば，分数 $x_r/\sum_1^n x_\cdot$ の i 個が x をこえる確率は問題**15**の (b) と同一である．同様に最大の分数が x をこえる確率は (a) に等しい．これらの分布は Fisher (1929) によって調和解析の有意性検定に使われている．

17　非心 f 分布と非心ベータ分布　$\chi_1^2\sim\chi^2(k,\lambda)$ [非心カイ2乗分布 (3b. 2. 2)] と $\chi_2^2\sim\chi^2(m)$ は独立であるとする．$f=\chi_1^2/\chi_2^2$ と $g=\chi_1^2/(\chi_1^2+\chi_2^2)$ を定義する．$\lambda=0$ のとき f は中心カイ2乗の比の分布 (3b. 1. 10) に従い，g はベータ分布 (3b. 1. 12) に従うことは示した．$\lambda\neq 0$ のとき χ_1^2 と χ_2^2 の同時密度関数は

$$e^{-\lambda/2}\sum\left(\frac{\lambda}{2}\right)^r\frac{1}{r!}\,G\left(\chi_1^2\middle|\frac{1}{2},\ \frac{k}{2}+r\right)d\chi_1^2 G\left(\chi_2^2\middle|\frac{1}{2},\ \frac{m}{2}\right)d\chi_2^2$$

である．

17. 1　各項に (3b. 1. 7) の結果を適用すると $f=\chi_1^2/\chi_2^2$ の密度関数

$$e^{-\lambda/2}\sum_{r=0}^{\infty}\left(\frac{\lambda}{2}\right)^r\frac{1}{r!}\frac{\Gamma\left(\frac{k+m}{2}+r\right)}{\Gamma\left(\frac{k}{2}+r\right)\Gamma\left(\frac{m}{2}\right)}\frac{f^{r-1+k/2}}{(1+f)^{r+(k+m)/2}}$$

が得られる．これは非心 f (カイ2乗の比) 分布と呼ばれ，$f(k,m,\lambda)$ であらわす．また超幾何関数 ${}_1F_1$ を用いると

$$e^{-\lambda/2}\frac{\Gamma\left(\frac{k+m}{2}\right)}{\Gamma\left(\frac{k}{2}\right)\Gamma\left(\frac{m}{2}\right)}\frac{f^{(k/2)-1}}{(1+f)^{(k+m)/2}}\,{}_1F_1\left(\frac{k+m}{2},\ \frac{k}{2},\ \frac{\lambda}{2}\frac{f}{1+f}\right)=f(f|k,m,\lambda)$$

という形に書ける．$F=mf/k$ を代入すると非心 F 分布 $F(k,m,\lambda)$ が得られる．

17. 2　変数変換により $g=\chi_1^2/(\chi_1^2+\chi_2^2)$ の密度関数

$$e^{-\lambda/2}\sum_{r=0}^{\infty}\left(\frac{\lambda}{2}\right)^r\frac{1}{r!}\,B\left(g\middle|\frac{k}{2}+r,\frac{m}{2}\right)$$

を得る．これは非心ベータ分布と呼ばれ $B(k,m,\lambda)$ であらわす．これはまた

$$e^{-\lambda/2}B\left(g\middle|\frac{k}{2},\ \frac{m}{2}\right){}_1F_1\left(\frac{k+m}{2},\ \frac{k}{2},\ \frac{\lambda g}{2}\right)=B(g|k,m,\lambda)$$

という別の形にも書ける．

18　$p(x)$ を確率変数の密度関数とし，エントロピーを $-\int p\log p\,dx$ で定義する．

18. 1　確率変数が有限な区間で有界，すなわち $a\leqslant x\leqslant b$ であるとき，エントロピーは a と b の間の一様分布にたいして最大となることを示せ．

18. 2　X を非負の値のみをとり，$E(X)=\mu>0$ (一定値) となるような確率変数とする．密度関数が $\mu^{-1}\exp(-x/\mu)$ という形，すなわち指数分布のときにエントロピーは最大になることを示せ．[ヒント：(3a. 1. 13) から (3a. 1. 15)，また (3a. 6. 1) から (3a. 6. 5) で述べた方法に従え．]

19　ベッセル関数分布　関数

$$I_\rho(x)=\sum_{k=0}^{\infty}\frac{1}{k!\Gamma(\rho+k+1)}\left(\frac{x}{2}\right)^{2k+\rho}$$

は変形ベッセル関数と呼ばれる ((3a. 7. 11) 参照)．

ガンマ分布 $G(\alpha,\rho+k+1)$ に従う確率変数 X を考え，k としてポアソン分布 $P(\lambda)$ に従う整数値の確率変数をとる．k について加えると得られる X の複合分布の密度関数は

$$\alpha\lambda^{-\rho}e^{-\lambda-\alpha x}\sum_0^{\infty}\frac{(x\alpha\lambda)^{\rho+k}}{k!\Gamma(\rho+k+1)}=\alpha e^{-\lambda-\alpha x}(x\alpha/\lambda)^{\rho/2}I_\rho(2\sqrt{\alpha\lambda x})$$

である．

20.　U と V が独立で V が正規分布に従うとき確率変数 X が $U+V$ であらわされるならば X は正規成分をもつという．このとき，一意な最大分解 $X=U+V$ が存在することを示せ．ただし U と V は独立で V は正規分布に従い U は正規成分を持たない．

X がベクトルの確率変数のとき，最大分解が存在するが一意でないことを示す例をあげよ．

21　正規分布の特徴づけ

21. 1　X_1, X_2 を独立な確率変数とする．X_1+X_2 と X_1-X_2 もまた独立ならば X_1 と X_2 は正規分布に従うことを示せ．

21. 2　ダルモア＝スキトヴィッチの定理　$X_1,\cdots,X_n$ を独立な確率変数とする．$a_1X_1+\cdots+a_nX_n$ と $b_1X_1+\cdots+b_nX_n$ が互いに独立な分布に従うならば X_i は $a_ib_i\neq 0$ のとき正規分布に従わなければならない [Darmois (1951), Skitovic (1954)].

21. 3　X_1, X_2 が $E(X_1)=0$ となるような互いに独立に同一の分布に従う確率変数ならば $E(a_1X_1+a_2X_2|b_1X_1+b_2X_2)=0$ は $a_1b_1+a_2b_2=0$ のとき X_1 が正規分布に従うことを意味している．

21. 4　$X_1,\cdots,X_n$ が $E(X_1)=0$ となるような互いに独立に同一の分布に従う確率変数ならば $E(\bar{X}|X_1-\bar{X})=0$ は X_1 が正規分布に従うことを意味している．(この証明や，その他のいろいろな結果については Kagan, Linnik and Rao(1972)による本を参照せよ．また Pathak and Pillai(1968) を参照せよ．)

参考文献

Bock, R. D. (1969), Estimating multinomial response relation, *Contributions to Statistics and Probability: Essays in Memory of S. N. Roy*, 111–132.

Cochran, W. G. (1934), The distribution of quadratic forms in a normal system, with applications to the analysis of covariance, *Proc. Camb. Phil. Soc.* **30**, 178–191.

Craig, A. T. (1943), Note on the independence of certain quadratic forms, *Ann. Math. Statist.* **14**, 195–197.

Cramer, H. (1937), *Random Variables and Probability Distributions*, Cambridge Tracts in Mathematics and Math. Physics **36**, Cambridge University Press.

Darmois, G. (1951), Sur une propriete caracteristique de la loi de probabilite de Laplace, *C. R.* **232**, 1999–2000.

Dempster, A. P. (1971), An overview of multivariate data analysis, *J. Multivariate Analysis* **1**, 316–346.

Fisher, R. A. (1915), Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population. *Biometrika* **10**, 507-521.

Fisher, R. A. (1924), On a distribution yielding the error function of several well-known statistics. *Proc. Int. Math. Congress*, 805-813.

Fisher, R. A. (1929), Tests of significance in harmonic analysis, *Proc. Roy. Soc.* **A125**, 54-59.

Fisher, R. A. (1953), Dispersion on a sphere, *Proc. Roy. Soc.* **A217**, 295-305.

Graybill, F. A. (1961), *Introduction to Linear Statistical Models*, McGraw-Hill, New York.

Hogg, R. V. and A. T. Craig (1958), On the decomposition of certain χ^2 variables, *Ann. Math. Statist.* **29**, 608-610.

Joos, Georg (1934), *Theoretical Physics*, Haffner, New York. (Earlier German edition, 1894.)

Linnik, Yu. V. (1952), Linear statistics and the normal law, *Dokl. Akad. Nauk SSSR* **83**, 353-355.

Mardia, K. V. (1972), *Statistics of Directional Data*, Academic Press, New York.

Mayer, J. W. and M. G. Mayer (1940), *Statistical Mechanics*, Wiley, New York.

Ogasawara, T. and M. Takahashi (1951), Independence of quadratic forms in normal system, *J. Sci. Hiroshima University* **15**, 1-9.

Ogawa, J. (1949), On the independence of linear and quadratic forms of a random sample from a normal population, *Ann. Inst. Math. Stat.* **1**, 83-108.

Pathak, P. K. and R. N. Pillai (1968), On a characterization of the normal law, *Sankhyā A* **30**, 141-144.

Polya, G. (1923), Herleitung des Gausschen Gesetzes aus einer Funktionalgleichung, *Math. Zeitschrift* **18**, 96-108.

Rao, J. S. and S. Sengupta (1972), Mathematical techniques for paleocurrent analysis, *J. Int. Assoc. Math., Geology* **4**, 235-248.

Sakamoto, H. (1944), On independence of statistics(in Japanese), *Res. Memoirs Instit. Stat. Math.* **1**, 1-25.

Skitovic, V. P. (1954), Linear combinations of independent random variables and the normal distribution law, *Izv. Akad. Nauk SSSR Ser. Mat.* **18**, 185-200.

Stevens, W. L. (1939), Solution to a geometrical problem in probability, *Annals of Eugenics* **9**, 315-320.

"Student" (W. S. Gosset) (1908), The probable error of a mean, *Biometrika* **6**, 1-25.

Watson, G. S. (1956), Analysis of dispersion on a sphere, *Monthly Notices of the Royal Astronomical Society* **7**, 153-159.

Watson, G. S. and E. J. Williams (1956), On the construction of significance tests on the circle and the sphere, *Biometrika* **43**, 344-352.

第4章　最小2乗法と分散分析の理論

まえがき　この章は，観測値の期待値にたいしては線形模型を，その分散共分散にたいしてはある特定の構造を想定したときの統計的推測を取り扱う．

最小2乗法の理論は，線形模型における母数の推定に関係がある．この理論の創設は Gauss (1809)，もっと最近では Markoff (1900) であった．最近の研究者によりある種の改良がなされた (Aitken (1935), Bose (1950—1951), Neyman and David (1938), Parzen (1961), Rao (1945b, c, 1946b, 1962f)．しかしながら，それの一般性の問題，特に観測値が特異な分散共分散行列をもつときが未解決の問題として残っている．後者の場合の特別な解が Goldman and Zelen (1964), Mitra and Rao (1968f), Zyskind and Martin (1969) などによって与えられている．

分散共分散行列の非特異性または特異性に関係しない，完全に一般的な形での問題の統一的接近が Rao (1971e, 1972d, 1973a) と Rao and Mitra (1971h) により与えられている．これらはこの章の **4i** 節に述べられている．この接近においては，主要な道具として，**1b.5** と **1c.5** で展開された行列の一般逆行列（g 逆行列）が用いられる．

4f では，2元分類の釣り合いのとれたデータにたいして，線形模型における級内相関の推定を取り扱うけれども，分散成分の推定の一般問題は，著者 (Rao (1970a, 1971c, 1971d, 1972b)) により最近発展させられた MINQUE（最小ノルム2次不偏推定）の理論と呼ばれているものを用いて **4j** で考える．

回帰と最小2乗予測に関する **4g** 節は多次元分布の一般的な形で取り扱われ，最小2乗法の理論と独立に研究することができる．重相関係数，偏相関係数や相関比のような関連性の尺度が，実際問題と特に関係づけた新しい観点から発展させられている．

ある種の母数の推定において最大の精度が得られるように，観測値を生成するような観測の計画について，簡単な議論が与えられる．

4a　最小2乗法の理論（線形推定論）

4a.1　ガウス＝マルコフ模型 $(\mathbf{Y}, \mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I})$

互いに無相関である観測値 $(y_1, \cdots, y_n)$ が，

$$\left.\begin{aligned} E(y_i) &= x_{i1}\beta_1 + \cdots + x_{im}\beta_m \\ V(y_i) &= \sigma^2 \end{aligned}\right\} \quad i = 1, 2, \cdots, n, \tag{4a. 1. 1}$$

であるとする．ここで $(\beta_1, \cdots, \beta_m)$ と σ^2 は未知の母数であり，$(x_{ij}) = \mathbf{X}$ は既知の係数の行列である．変量 y_i，母数 β_j の列ベクトルをそれぞれ $\mathbf{Y}, \boldsymbol{\beta}$ であらわすと，(4a. 1. 1) は行列記号を用いて，

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon},\ E(\boldsymbol{\varepsilon}) = \mathbf{0},\ D(\boldsymbol{\varepsilon}) = \sigma^2\mathbf{I} \Rightarrow E(\mathbf{Y}) = \mathbf{X}\boldsymbol{\beta},\ D(\mathbf{Y}) = \sigma^2\mathbf{I} \tag{4a. 1. 2}$$

と書くことができる．ここで D は分散共分散をあらわし，$\mathbf{I}$ は n 次の単位行列をあらわす．問題は，観測値 y_i をもとにして未知の母数 β_j を推定することである．

Aitken (1935) によって考えられた (4a. 1. 2) より少し一般的な模型は，

$$E(\mathbf{Y}) = \mathbf{X}\boldsymbol{\beta}, \qquad D(\mathbf{Y}) = \sigma^2\mathbf{G}, \qquad |\mathbf{G}| \neq 0 \tag{4a. 1. 3}$$

であって，観測値の間に相関をいれている．(4a. 1. 3) において，$\mathbf{G}$ は既知の行列であり，$\boldsymbol{\beta}, \sigma^2$ は未知である．ここで $\mathbf{Z} = \mathbf{G}^{-\frac{1}{2}}\mathbf{Y}$ とおくと，

$$\begin{aligned} E(\mathbf{Z}) &= \mathbf{G}^{-\frac{1}{2}}\mathbf{X}\boldsymbol{\beta} = \mathbf{U}\boldsymbol{\beta}, \\ D(\mathbf{Z}) &= \sigma^2\mathbf{I} \end{aligned} \tag{4a. 1. 4}$$

であるから，模型 (4a. 1. 3) は模型 (4a. 1. 2) に帰着させられる．(4a. 1. 3) の1つの特別な場合は

$$E(\mathbf{Y}) = \mathbf{X}\boldsymbol{\beta}, \qquad D(\mathbf{Y}) = \boldsymbol{\Sigma}, \qquad |\boldsymbol{\Sigma}| \neq 0 \tag{4a. 1. 5}$$

である．ここで $\boldsymbol{\Sigma}$ は既知である．この場合には，変換 $\mathbf{Z} = \boldsymbol{\Sigma}^{-\frac{1}{2}}\mathbf{Y}$ を行うと

$$\begin{aligned} E(\mathbf{Z}) &= \boldsymbol{\Sigma}^{-\frac{1}{2}}\mathbf{X}\boldsymbol{\beta} = \mathbf{U}\boldsymbol{\beta} \\ D(\mathbf{Z}) &= \mathbf{I} \end{aligned} \tag{4a. 1. 6}$$

であるから，模型 (4a. 1. 5) は模型 (4a. 1. 2) に帰着させられる．したがってこれらすべての場合にたいして，模型 (4a. 1. 2) だけを詳細に考えておけばよい．模型 (4a. 1. 2) を $(\mathbf{Y}, \mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I})$ であらわす．

$(\mathbf{Y}, \mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{G})$ において，$\mathbf{G}$ が特異であってもよいという一般の場合には，(4a. 1. 4) の変換は適用不可能である．この場合は **4i** 節で考える．

次の記号と結果を使う．読者は，これらの結果を確かめた上，よく理解しておいてもらいたい．

(a)　$\mathbf{L}'\mathbf{Y}, \mathbf{M}'\mathbf{Y}, \cdots$ は $\mathbf{Y}$ の線形関数をあらわし，$\mathbf{P}'\boldsymbol{\beta}, \mathbf{R}'\boldsymbol{\beta}, \cdots$ は $\boldsymbol{\beta}$ の線形関数をあらわす．ここで $\mathbf{L}, \mathbf{M}, \cdots, \mathbf{P}, \mathbf{R}, \cdots$ は列ベクトルをあらわす．

(b)　$E(\mathbf{L}'\mathbf{Y}) = \mathbf{L}'E(\mathbf{Y}) = \mathbf{L}'\mathbf{X}\boldsymbol{\beta}$.

(c)　$V(\mathbf{L}'\mathbf{Y}) = \mathbf{L}'\boldsymbol{\Sigma}\mathbf{L}$,　ここで $\boldsymbol{\Sigma}$ は $\mathbf{Y}$ の分散共分散行列をあらわす

$= \sigma^2\mathbf{L}'\mathbf{L}$,　$\boldsymbol{\Sigma} = \sigma^2\mathbf{I}$ の場合

$\mathrm{cov}(\mathbf{L}'\mathbf{Y}, \mathbf{M}'\mathbf{Y}) = \mathbf{L}'\boldsymbol{\Sigma}\mathbf{M}$,　一般の $\boldsymbol{\Sigma}$ にたいして

$= \sigma^2\mathbf{L}'\mathbf{M}$,　$\boldsymbol{\Sigma} = \sigma^2\mathbf{I}$ の場合.

(d)　もっと一般には，$\mathbf{B}$ が $n \times k$ の行列のときには $\mathbf{B}'\mathbf{Y}$ は $\mathbf{Y}$ の k 個の線形関数をあらわす．次のことが容易に証明される．$E(\mathbf{B}'\mathbf{Y}) = \mathbf{B}'\mathbf{X}\boldsymbol{\beta}$, $D(\mathbf{B}'\mathbf{Y}) = \mathbf{B}'\boldsymbol{\Sigma}\mathbf{B}$, $\boldsymbol{\Sigma} = \sigma^2\mathbf{I}$ のときには後者の式は $\sigma^2\mathbf{B}'\mathbf{B}$ となる．

(e)　2次形式の期待値：

$$\begin{aligned} E(\mathbf{Y}'\mathbf{G}\mathbf{Y}) &= \mathrm{trace}\,\mathbf{G}\boldsymbol{\Sigma} + E(\mathbf{Y}')\mathbf{G}\,E(\mathbf{Y}), \qquad \text{ここで } \boldsymbol{\Sigma} = D(\mathbf{Y}) \\ &= \sigma^2\,\mathrm{trace}\,\mathbf{G} + E(\mathbf{Y}')\mathbf{G}\,E(\mathbf{Y}), \qquad \boldsymbol{\Sigma} = \sigma^2\mathbf{I} \text{ の場合} \end{aligned} \tag{4a. 1. 7}$$

(f)　$\mathbf{A}$ がベキ等行列ならば，$\mathrm{rank}\,\mathbf{A} = \mathrm{trace}\,\mathbf{A}$

(g)　行列 $\mathbf{S}$ が与えられたとき，性質

$$\mathbf{S}\mathbf{S}^-\mathbf{S} = \mathbf{S}, \qquad \mathrm{trace}\,\mathbf{S}^-\mathbf{S} = \mathrm{rank}\,\mathbf{S} \tag{4a. 1. 8}$$

をもつ g 逆行列と呼ばれる行列 $\mathbf{S}^-$ が存在する (**1b. 5** を参照)．方程式 $\mathbf{S}\boldsymbol{\beta} = \mathbf{Q}$ が解をもつとき，これの一般解は $\mathbf{S}^-\mathbf{Q} + (\mathbf{I} - \mathbf{S}^-\mathbf{S})\mathbf{Z}$ である．ここで $\mathbf{Z}$ は任意である．$\mathbf{S}$ が正方行列で正則である場合には，$\mathbf{S}^- = \mathbf{S}^{-1}$，つまり $\mathbf{S}^-$ は真の逆行列である．

(h)　$\mathbf{A}'$ の列ベクトルによってつくられる線形部分空間と $\mathbf{A}'\mathbf{A}$ の列ベクトルによってつくられる線形部分空間とは同じである．つまり $\mathscr{M}(\mathbf{A}') = \mathscr{M}(\mathbf{A}'\mathbf{A})$ である．このことから，$\mathbf{P} = \mathbf{A}'\mathbf{L}$ であるようなベクトル $\mathbf{L}$ が存在するならば，$\mathbf{P} = \mathbf{A}'\mathbf{A}\boldsymbol{\lambda}$ であるようなベクトル $\boldsymbol{\lambda}$ が存在することになる．任意の $\mathbf{L}$ にたいして $\mathbf{A}'\mathbf{L} \in \mathscr{M}(\mathbf{A}') = \mathscr{M}(\mathbf{A}'\mathbf{A})$ であることを確かめよ．

4a. 2　正規方程式と最小2乗 (least square, l. s.) 推定量

$\boldsymbol{\beta}$ に関する方程式

$$\mathbf{X}'\mathbf{X}\boldsymbol{\beta} = \mathbf{X}'\mathbf{Y} \tag{4a. 2. 1}$$

は正規方程式と呼ばれる．これは，観測値とそれの期待値との差の2乗和

$$(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}) = \sum (y_i - x_{i1}\beta_1 - \cdots - x_{im}\beta_m)^2 \tag{4a. 2. 2}$$

を（最小にするために）微分することにより得られる（第1章の終りの問題 **13. 2** を参照）．観測方程式 $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta}$ は一般には解をもたない．しかし正規方程式 (4a. 2. 1) は常に解をもつ．なぜならば，$\mathbf{X}'\mathbf{Y} \in \mathscr{M}(\mathbf{X}'\mathbf{X})$ であるから (4a. 2. 1) が成り立つような $\boldsymbol{\beta}$ は存在する．$\hat{\boldsymbol{\beta}}$ を (4a. 2. 1) の任意の解とする．そうすると

$$\begin{aligned} (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}) &= [\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}} + \mathbf{X}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})]'[\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}} + \mathbf{X}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})] \\ &= (\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})'(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}) + (\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})'\mathbf{X}'\mathbf{X}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \\ &\geqslant (\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})'(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}), \end{aligned} \tag{4a. 2. 3}$$

このことは，$(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})$ の最小値が

$$(\mathbf{Y}-\mathbf{X}\hat{\beta})'(\mathbf{Y}-\mathbf{X}\hat{\beta})$$

であって，それは $\beta=\hat{\beta}$ で到達され，(4a. 2. 1) のすべての解にたいして一意であるということを示す．

母数の関数 $\mathbf{P}'\beta$ の最小2乗推定量を $\mathbf{P}'\hat{\beta}$ と定義する．ここで β は正規方程式 (4a. 2. 1) の任意の解である．推定量 $\mathbf{P}'\hat{\beta}$ はどのような性質をもち，またどのような意味において $\mathbf{P}'\beta$ のよい推定量であるか？

（i）$\mathbf{P}'\hat{\beta}$ が $\mathbf{Y}$ に関して線形でありかつ $\mathbf{P}'\beta$ の不偏推定量であるための必要十分条件は，正規方程式のすべての解 β にたいして $\mathbf{P}'\hat{\beta}$ が一意であることである．この条件は，$\mathbf{P}\in\mathscr{M}(\mathbf{X}')$，(これは $\mathbf{P}\in\mathscr{M}(\mathbf{X}'\mathbf{X})$ と同じである)，と同値である．また，この条件は，期待値が $\mathbf{P}'\beta$ であるような $\mathbf{Y}$ の線形関数が存在するということとも同値である．

未知母数 β_j の数 m や，行列 $\mathbf{X}$ の階数については何の仮定もしない．

$\mathbf{P}\in\mathscr{M}(\mathbf{X}'\mathbf{X})$ とする．すなわち，あるベクトル λ にたいして $\mathbf{P}=\mathbf{X}'\mathbf{X}\lambda$ とする．そうすると任意の解 $\hat{\beta}$ にたいして

$$\mathbf{X}'\mathbf{X}\hat{\beta}=\mathbf{X}'\mathbf{Y}\ \text{であるから，}\ \mathbf{P}'\hat{\beta}=\lambda'\mathbf{X}'\mathbf{X}\hat{\beta}=\lambda'\mathbf{X}'\mathbf{Y}. \tag{4a. 2. 4}$$

したがって，$\mathbf{P}'\hat{\beta}$ は一意であって $\mathbf{Y}$ に関して線形である．また，$E(\lambda'\mathbf{X}'\mathbf{Y})=\lambda'\mathbf{X}'\mathbf{X}\beta=\mathbf{P}'\beta$，すなわち $\mathbf{P}'\hat{\beta}$ は $\mathbf{P}'\beta$ の不偏推定量である．さらに，λ' は一意ではないかも知れないが $\lambda'\mathbf{X}'$ は一意である．

$\mathbf{P}'\hat{\beta}$ は一意であるとする．$\mathbf{X}'\mathbf{X}\lambda=0$ であるような任意のベクトル λ に対応して，(4a. 2. 1) の解 $\bar{\beta}=\hat{\beta}-\lambda$ がある．そして，$\mathbf{X}'\mathbf{X}\lambda=0$ であるようなすべての λ にたいして，$0=\mathbf{P}'(\hat{\beta}-\bar{\beta})=\mathbf{P}'\lambda$．したがって $\mathbf{P}\in\mathscr{M}(\mathbf{X}'\mathbf{X})$．

$\mathbf{L}'\mathbf{Y}$ を $E(\mathbf{L}'\mathbf{Y})=\mathbf{P}'\beta$ であるものとする．そうすると $\mathbf{L}'\mathbf{X}\beta=\mathbf{P}'\beta$ または $\mathbf{L}'\mathbf{X}=\mathbf{P}'$ が得られる．このことは $\mathbf{P}\in\mathscr{M}(\mathbf{X}')$ を示す．

(i) における条件のどれかが成立するとき，$\mathbf{P}'\beta$ は推定可能であると呼ぶ．

（ii）$\mathbf{P}'\beta$ が推定可能であるならば，$\mathbf{P}'\hat{\beta}$ は，$\mathbf{P}'\beta$ の線形不偏推定量 LUE (linear unbiased estimator) のなかで最小分散をもつ．

$\mathbf{P}'\beta$ は推定可能であるから $\mathbf{P}\in\mathscr{M}(\mathbf{X}'\mathbf{X})$，すなわち，ある λ にたいして $\mathbf{P}=\mathbf{X}'\mathbf{X}\lambda$．したがって

$$\mathbf{P}'\hat{\beta}=\lambda'\mathbf{X}'\mathbf{X}\hat{\beta}=\lambda'\mathbf{X}'\mathbf{Y}. \tag{4a. 2. 5}$$

$\mathbf{L}'\mathbf{Y}$ を $\mathbf{P}'\beta$ の任意の不偏推定量とすると，$\mathbf{L}'\mathbf{X}=\mathbf{P}'$．

$$\begin{aligned}V(\mathbf{L}'\mathbf{Y})&=V(\mathbf{L}'\mathbf{Y}-\lambda'\mathbf{X}'\mathbf{Y}+\lambda'\mathbf{X}'\mathbf{Y})\\&=V(\mathbf{L}'\mathbf{Y}-\lambda'\mathbf{X}'\mathbf{Y})+V(\lambda'\mathbf{X}'\mathbf{Y})\\&=V(\mathbf{L}'\mathbf{Y}-\lambda'\mathbf{X}'\mathbf{Y})+V(\mathbf{P}'\hat{\beta})\end{aligned} \tag{4a. 2. 6}$$

である．ただし，(4a. 2. 6) において，$(\mathbf{L}'-\lambda'\mathbf{X}')\mathbf{Y}$ と $\lambda'\mathbf{X}'\mathbf{Y}$ との共分散は

$$(\mathbf{L}'-\lambda'\mathbf{X}')\mathbf{X}\lambda\sigma^2=(\mathbf{L}'\mathbf{X}-\lambda'\mathbf{X}'\mathbf{X})\lambda\sigma^2=(\mathbf{P}'-\mathbf{P}')\lambda\sigma^2=0$$

である．(4a. 2. 6) から $V(\mathbf{L}'\mathbf{Y})\geqq V(\mathbf{P}'\hat{\beta})$．$\mathbf{P}'\hat{\beta}$ を $\mathbf{P}'\beta$ の BLUE（最良の LUE）と呼ぶ．

（iii）母数の線形関数がすべて推定可能であるための必要十分条件は，$R(\mathbf{X})=m$（未知母数 β_j の数）である．ここで $R(\mathbf{X})$ は $\mathbf{X}$ の階数をあらわす．

$R(\mathbf{X})=m$ ならば $R(\mathbf{X}'\mathbf{X})=m$ となり正規方程式は一意の解をもつ．したがって $\mathbf{P}'\hat{\beta}$ は任意に与えられた $\mathbf{P}$ にたいして一意である．

$\mathbf{P}'\beta$ が推定可能であるならば $\mathbf{P}\in\mathscr{M}(\mathbf{X}'\mathbf{X})$．もし $\mathbf{P}$ が任意であるならば $R(\mathbf{X}'\mathbf{X})$ は最大階数でなければいけない．よって逆が示された．

$R(\mathbf{X})<m$ ならば，いくつかの母数の関数は不偏推定量をもたず，ここでの理論構成の立場からは，このような母数の関数については何の推測もできない．このような母数の関数は推定不可能である，または交絡されていると呼ばれる．

結果 (i) と (ii) は，母数の関数にたいして，最小分散をもつ線形不偏推定量を探すならば，ガウスの l. s. 推定量が答であるということを示している．推定可能の概念はボースによる．

（iv）$\mathbf{C}$ を $\mathbf{X}'\mathbf{X}$ の一般逆行列とし $\mathbf{H}=\mathbf{C}\mathbf{X}'\mathbf{X}$ とする．$\mathbf{P}'\beta$ が推定可能であるための必要十分条件は $\mathbf{P}'(\mathbf{I}-\mathbf{H})=0$ である．

$\mathbf{X}'\mathbf{X}\beta=\mathbf{X}'\mathbf{Y}$ の一般解は，$\mathbf{Z}$ を任意として $\mathbf{C}\mathbf{X}'\mathbf{Y}+(\mathbf{I}-\mathbf{H})\mathbf{Z}$ であることが示されている [(vi), 1b. 5]．推定可能にたいする (i) で与えた必要十分条件，すなわち，すべての解 $\hat{\beta}$ にたいして $\mathbf{P}'\hat{\beta}$ が一意であるを用いると，すべての $\mathbf{Z}$ にたいして $\mathbf{P}'(\mathbf{I}-\mathbf{H})\mathbf{Z}=0\Leftrightarrow\mathbf{P}'(\mathbf{I}-\mathbf{H})=0$ であることがわかる．

4a. 3　一般逆行列と正規方程式の解

正規方程式 $\mathbf{X}'\mathbf{X}\beta=\mathbf{X}'\mathbf{Y}$ が解をもつことはすでに示されている．したがって $\mathbf{X}'\mathbf{X}$ の1つの一般逆行列 $\mathbf{C}$ がわかれば，特解は $\mathbf{C}\mathbf{X}'\mathbf{Y}$ によって与えられる．実際には $\mathbf{C}$ の計算は適当な行列の真の逆行列を求める計算に帰着させられる．$\mathbf{C}$ を計算するためのいくつかの方法を考えよう．

1.　m を未知母数の個数とするとき，もし $\operatorname{rank}\mathbf{X}=m$ ならば，$R(\mathbf{X}'\mathbf{X})=m$ で $\mathbf{C}=(\mathbf{X}'\mathbf{X})^{-1}$ は真の逆行列である．

2.　$R(\mathbf{X})=r<m$ とする．大きさ $(m-r)\times m$ の行列 $\mathbf{H}$ を $R(\mathbf{X}'\vdots\mathbf{H}')=m$ であるように定めることができるとする．すなわち，階数の不足を補うために $\mathbf{X}$ にいくつかの行をつけ加える．そうすると $R(\mathbf{X}'\mathbf{X}+\mathbf{H}'\mathbf{H})=m$ となり，真の逆行列 $(\mathbf{X}'\mathbf{X}+\mathbf{H}'\mathbf{H})^{-1}$ が存在する．$\mathbf{C}=(\mathbf{X}'\mathbf{X}+\mathbf{H}'\mathbf{H})^{-1}$ は $\mathbf{X}'\mathbf{X}$ の一般逆行列すなわち，$\mathbf{X}'\mathbf{X}\mathbf{C}\mathbf{X}'\mathbf{X}=\mathbf{X}'\mathbf{X}$ であることが容易に示される (32ページの問題5を参照)．

3.　(2) におけるような $\mathbf{H}$ がきめられたとき，一般逆行列を計算する他の方法がある．

$$\left(\begin{array}{c:c}\mathbf{X'X} & \mathbf{H'}\\ \hdashline \mathbf{H} & \mathbf{0}\end{array}\right)$$

の階数は $2m-r$ であることに注意せよ．したがって，真の逆行列

$$\left(\begin{array}{c:c}\mathbf{X'X} & \mathbf{H'}\\ \hdashline \mathbf{H} & \mathbf{0}\end{array}\right)^{-1}=\left(\begin{array}{c:c}\mathbf{C}_1 & \mathbf{C}_2'\\ \hdashline \mathbf{C}_2 & \mathbf{C}_4\end{array}\right) \tag{4a. 3. 1}$$

は存在する．逆行列 (4a. 3. 1) の分割形において現われる行列 $\mathbf{C}_1$ は確かに $\mathbf{X'X}$ の一般逆行列である，すなわち，方程式 $\mathbf{X'XC_1X'X}=\mathbf{X'X}$ を満足する．

4.　(3) の1つの修正は

$$(\mathbf{C}_1' \,\vdots\, \mathbf{C}_2')\left(\begin{array}{c}\mathbf{X'X}\\ \hdashline \mathbf{H}\end{array}\right)=\mathbf{I}$$

であるような行列 $\mathbf{C}_1', \mathbf{C}_2'$ を求めることである．そうすると $\mathbf{C}_1$ は $\mathbf{X'X}$ の一般逆行列である．多くの統計の問題においては，行列 $\mathbf{C}_1', \mathbf{C}_2'$ は視察によって書き下ろすことができる．

5.　$\mathbf{X'X}$ における1次従属な行をみいだすことができるとする．これらの行とこれらに対応する列とを除くと階数 r の $r\times r$ 行列が得られ，これは真の逆行列をもつ．この逆行列に，従属な行と列とが取り除かれたところに，ゼロからなる行と列とを挿入して行列を大きくする．こうして得られた行列は一般逆行列である．実際には，$\mathbf{X'X}$ においてどの行が1次従属かがわかることがあるので，これは有用な方法である．

6.　$\mathbf{X'X}$ のゼロでない固有値を $\lambda_1,\cdots,\lambda_r$ とし，対応する固有ベクトルを $\mathbf{P}_1,\cdots,\mathbf{P}_r$ とする．そうすると

$$\mathbf{X'X}=\lambda_1\mathbf{P}_1\mathbf{P}_1'+\cdots+\lambda_r\mathbf{P}_r\mathbf{P}_r'.$$

$$\mathbf{C}=(\mathbf{X'X})^-=\frac{1}{\lambda_1}\mathbf{P}_1\mathbf{P}_1'+\cdots+\frac{1}{\lambda_r}\mathbf{P}_r\mathbf{P}_r'$$

が一般逆行列であることは容易に確かめられる．

一般逆行列の計算に関する一般の議論については Rao and Mitra (1971B3) を参照せよ．

4a. 4　l. s. 推定量の分散と共分散

ここでは正規方程式 $\mathbf{X'X\beta}=\mathbf{X'Y}$ を $\mathbf{S\beta}=\mathbf{Q}$ であらわしておくのが便利である．ただし $\mathbf{S}=\mathbf{X'X}$, $\mathbf{Q}=\mathbf{X'Y}$. 正規方程式の本質的な性質は $D(\mathbf{Q})=\sigma^2\mathbf{S}$ である．$\mathbf{S}$ の一般逆行列を $\mathbf{C}$, すなわち $\mathbf{SCS}=\mathbf{S}$ とすると，正規方程式の1つの解の具体的な表現として $\hat{\boldsymbol{\beta}}=\mathbf{CQ}$ が得られる．

(i)　推定可能な関数 $\mathbf{P'\beta}$, $\mathbf{R'\beta}$ の l. s. 推定量を $\mathbf{P'}\hat{\boldsymbol{\beta}}$, $\mathbf{R'}\hat{\boldsymbol{\beta}}$ とし，$\mathbf{S}$ の任意の一般逆行列を $\mathbf{C}$ とする．そうすると

$$V(\mathbf{P'}\hat{\boldsymbol{\beta}})=\sigma^2\mathbf{P'CP},\qquad \mathrm{cov}(\mathbf{P'}\hat{\boldsymbol{\beta}},\mathbf{R'}\hat{\boldsymbol{\beta}})=\sigma^2\mathbf{P'CR} \tag{4a. 4. 1}$$

である．したがって，公式 (4a. 4. 1) を推定可能な母数の関数に適用する限り，$\sigma^2\mathbf{C}$ を $\hat{\boldsymbol{\beta}}$ の分散共分散行列と形式的に考えてよい．

$\mathbf{P'\beta}$ は推定可能であるから，$\mathbf{P}=\mathbf{X'X\lambda}=\mathbf{S\lambda}$. さて

$$\begin{aligned}V(\mathbf{P'}\hat{\boldsymbol{\beta}})&=V(\mathbf{P'CQ})=\mathbf{P'C}D(\mathbf{Q})\mathbf{C'P}\\&=\sigma^2\mathbf{P'CSC'P}=\sigma^2\boldsymbol{\lambda}'\mathbf{SCSC'P}=\sigma^2\mathbf{P'CP}.\end{aligned}$$

共分散に関する結果も同様にして証明される．

$\mathbf{C}=(C_{ij})$ が真の逆行列である場合にはすべての母数の関数は推定可能である，すなわち $\hat{\beta}_i$ は β_i の l. s. 推定量であり

$$V(\hat{\beta}_i)=\sigma^2C_{ii},\qquad \mathrm{cov}(\hat{\beta}_i,\hat{\beta}_j)=\sigma^2C_{ij}. \tag{4a. 4. 2}$$

表現 (4a. 4. 2) は，$\mathbf{C}$ が真の逆行列でないときには無意味であるが，それでも，β_i に関する線形関数が推定可能であるとき，これに対応する $\hat{\beta}_i$ の線形関数の分散を計算するのに用いることができる．

l. s. 推定量やそれの分散・共分散を計算するのに，前もって $\mathbf{C}$ を求めておくことは必ずしも常に必要ではない．正規方程式 $\mathbf{S\beta}=\mathbf{Q}$ にあるベクトル $\boldsymbol{\lambda}$ をかけ，$\boldsymbol{\lambda}'\mathbf{S\beta}=\boldsymbol{\lambda}'\mathbf{Q}$ が $\mathbf{P'\beta}=\boldsymbol{\lambda}'\mathbf{Q}$ になったとする．そうすると $\boldsymbol{\lambda}'\mathbf{Q}$ は $\mathbf{P'\beta}$ の l. s. 推定量であり

$$V(\boldsymbol{\lambda}'\mathbf{Q})=\sigma^2\boldsymbol{\lambda}'\mathbf{S\lambda}=\sigma^2\boldsymbol{\lambda}'\mathbf{P}. \tag{4a. 4. 3}$$

したがって $\boldsymbol{\lambda}$ と $\mathbf{P}$ だけに関係する分散の表現が得られた．同様に $\boldsymbol{\lambda}'\mathbf{S}=\mathbf{P}'$, $\boldsymbol{\mu}'\mathbf{S}=\mathbf{R}'$ であるとすると，$\boldsymbol{\lambda}'\mathbf{Q}$ と $\boldsymbol{\mu}'\mathbf{Q}$ は $\mathbf{P'\beta}$ と $\mathbf{R'\beta}$ の l. s. 推定量であり

$$\mathrm{cov}(\boldsymbol{\lambda}'\mathbf{Q},\boldsymbol{\mu}'\mathbf{Q})=\sigma^2\boldsymbol{\lambda}'\mathbf{R}=\sigma^2\boldsymbol{\mu}'\mathbf{P}. \tag{4a. 4. 4}$$

多くの実際問題において，公式 (4a. 4. 3), (4a. 4. 4) は適用可能であり，一般逆行列の計算をさけることができる．

4a. 5　σ^2 の推定

(4a. 2. 3) で導入された2乗和の最小値 $R_0^2=(\mathbf{Y}-\mathbf{X}\hat{\boldsymbol{\beta}})'(\mathbf{Y}-\mathbf{X}\hat{\boldsymbol{\beta}})$ は次の異なった表現をもつ．証明については読者自ら確かめよ．

$$\begin{aligned}R_0^2&=\mathbf{Y'Y}-\mathbf{Y'X}\hat{\boldsymbol{\beta}}=\mathbf{Y'Y}-\hat{\boldsymbol{\beta}}'\mathbf{X'X}\hat{\boldsymbol{\beta}}\\&=\mathbf{Y'Y}-\mathbf{Q}'\hat{\boldsymbol{\beta}}=\mathbf{Y'Y}-\mathbf{Q'CQ}\\&=\mathbf{Y'Y}-\mathbf{Y'X}(\mathbf{X'X})^-\mathbf{X'Y}=\mathbf{Y'}(\mathbf{I}-\mathbf{X}(\mathbf{X'X})^-\mathbf{X'})\mathbf{Y}.\end{aligned} \tag{4a. 5. 1}$$

$|\mathbf{X'X}|\neq 0$ のときには，R_0^2 は次の比として表現される．

$$\begin{vmatrix}\mathbf{Y'Y} & \mathbf{Y'X}\\ \mathbf{X'Y} & \mathbf{X'X}\end{vmatrix}\div|\mathbf{X'X}|. \tag{4a. 5. 2}$$

ベクトル $\mathbf{Y}-\mathbf{X}\hat{\boldsymbol{\beta}}$ は残差ベクトルと呼ばれる．多くの問題において残差を調べることは

必要である．次の事実を証明しよう．

（i）$E(\mathbf{Y}-\mathbf{X}\hat{\boldsymbol{\beta}})=0$

$$E(\mathbf{Y}-\mathbf{X}\hat{\boldsymbol{\beta}})=E(\mathbf{Y})-E(\mathbf{X}\hat{\boldsymbol{\beta}})=\mathbf{X}\boldsymbol{\beta}-\mathbf{X}\boldsymbol{\beta}=0.\ \text{一般には}\ E(\hat{\boldsymbol{\beta}})\neq\boldsymbol{\beta}$$

であるから，$E(\mathbf{X}\hat{\boldsymbol{\beta}})=\mathbf{X}\boldsymbol{\beta}$ のステップは証明の必要があることに注意せよ．

（ii）$\mathbf{P}'\hat{\boldsymbol{\beta}}$ が推定可能な母数の関数 $\mathbf{P}'\boldsymbol{\beta}$ の推定量であるときには，$\mathrm{cov}(\mathbf{P}'\hat{\boldsymbol{\beta}},\ \mathbf{Y}-\mathbf{X}\hat{\boldsymbol{\beta}})=0$.

この命題は，もっと一般的な結果 $\mathrm{cov}(\mathbf{P}'\hat{\boldsymbol{\beta}},\mathbf{L}'\mathbf{Y})=0$ から得られる．ここで $\mathbf{L}'\mathbf{Y}$ は $E(\mathbf{L}'\mathbf{Y})=0$ であるような $\mathbf{Y}$ の線形関数である．$\mathbf{L}'\mathbf{Y}$ の条件から $\mathbf{L}'\mathbf{X}=\mathbf{0}$ が得られる．$\mathbf{P}=\mathbf{X}'\mathbf{X}\boldsymbol{\lambda}$ とすると $\mathbf{P}'\hat{\boldsymbol{\beta}}=\boldsymbol{\lambda}'\mathbf{X}'\mathbf{Y}$ であり

$$\mathrm{cov}(\mathbf{P}'\hat{\boldsymbol{\beta}},\mathbf{L}'\mathbf{Y})=\mathrm{cov}(\boldsymbol{\lambda}'\mathbf{X}'\mathbf{Y},\mathbf{L}'\mathbf{Y})=\sigma^2\mathbf{L}'\mathbf{X}\boldsymbol{\lambda}=0. \qquad (4a.5.3)$$

(i) より，$\mathbf{Y}-\mathbf{X}\hat{\boldsymbol{\beta}}$ は期待値がゼロである線形関数の集まりである．よって上の命題が得られる．

（iii）$\mathbf{Y}-\mathbf{X}\hat{\boldsymbol{\beta}}$ の分散共分散行列は分散共分散行列の差である，つまり

$$D(\mathbf{Y}-\mathbf{X}\hat{\boldsymbol{\beta}})=D(\mathbf{Y})-D(\mathbf{X}\hat{\boldsymbol{\beta}})=\sigma^2(\mathbf{I}-\mathbf{X}(\mathbf{X}'\mathbf{X})^-\mathbf{X}'). \qquad (4a.5.4)$$

(ii) より $\mathrm{cov}(\mathbf{Y}-\mathbf{X}\hat{\boldsymbol{\beta}},\mathbf{X}\hat{\boldsymbol{\beta}})=0$. したがって (4a.5.4) が得られる．

（iv）$E(R_0^2)=(n-r)\sigma^2$. したがって σ^2 の1つの不偏推定量は

$$\hat{\sigma}^2=\frac{R_0^2}{n-r} \qquad (4a.5.5)$$

である．ただし $r=R(\mathbf{X})$.

この結果を証明する最も簡単な方法は，$\mathbf{Y}$ から $\mathbf{Z}$ への直交変換［(i), 3b.5］と新しい変量での表現 (3b.5.4)，つまり $R_0^2=z_{r+1}^2+\cdots+z_n^2$ を用いることである．ここで $E(z_i)=0,\ V(z_i)=\sigma^2,\ i=r+1,\cdots,n$ である．したがって $E(R_0^2)=(n-r)\sigma^2$.

または，表現 (4a.5.1)

$$R_0^2=\mathbf{Y}'\mathbf{Y}-\mathbf{Y}'\mathbf{X}(\mathbf{X}'\mathbf{X})^-\mathbf{X}'\mathbf{Y}=\mathbf{Y}'(\mathbf{I}-\mathbf{X}(\mathbf{X}'\mathbf{X})^-\mathbf{X}')\mathbf{Y}$$

を用いてもよい．性質 $\mathbf{X}'\mathbf{X}(\mathbf{X}'\mathbf{X})^-\mathbf{X}'=\mathbf{X}'$（第3章，問題2.1）を用いると，$\mathbf{X}\boldsymbol{\beta}$ を入れても R_0^2 の値は変わらないことがわかる．

$$R_0^2=(\mathbf{Y}-\mathbf{X}\boldsymbol{\beta})'(\mathbf{I}-\mathbf{X}(\mathbf{X}'\mathbf{X})^-\mathbf{X}')(\mathbf{Y}-\mathbf{X}\boldsymbol{\beta})$$

$$E(R_0^2)=\sigma^2\mathrm{trace}(\mathbf{I}-\mathbf{X}(\mathbf{X}'\mathbf{X})^-\mathbf{X}')$$
$$=n\sigma^2-\sigma^2\mathrm{trace}(\mathbf{X}'\mathbf{X})^-(\mathbf{X}'\mathbf{X})=(n-r)\sigma^2$$

最後のところで，(4a.1.8) より $\mathrm{trace}(\mathbf{X}'\mathbf{X})^-(\mathbf{X}'\mathbf{X})=R(\mathbf{X}'\mathbf{X})=r$ であることを用いた．

（v）不偏性は別として，$R_0^2/(n-r)$ はどのような意味において σ^2 の最適な推定量であるか？ Hsu (1938) ト Rao (1952a, 1971d) は，σ^2 にたいして不偏であるような2次形式の推定量 $\mathbf{Y}'\mathbf{G}\mathbf{Y}$ を考え，この分散を最小にするという形でこの問題を調べた．スーは $V(\mathbf{Y}'\mathbf{G}\mathbf{Y})$ が未知母数 $\boldsymbol{\beta}$ に無関係であるという条件をつけ，ラオは $\mathbf{Y}'\mathbf{G}\mathbf{Y}$ が非負であるという条件をつけた．どちらの条件の下でも，（a）$\beta_2=3$，または（b）$\beta_2\neq3$ のときは，R_0^2 の展開における $y_1^2,\cdots,y_n^2$ の係数が同じである，のとき，$R_0^2/(n-r)$ は最小分散をもつことが示される．ここで β_2 は個々の y_i の分布の尖度にたいするピアソンの係数である．これ以上の結果は，β_2 の正確な値か y_i の分布を知らないと得られない．多くの場合において条件 (b) は満足されず，そのときには，最小2乗法に基づく σ^2 の推定量にたいして最小分散性を要求することはできない．

4a.6　最小2乗法の理論にたいする他の接近（幾何学的解）

$\mathbf{X}$ の列空間の上への射影作用素（行列，1c.4）を $\mathbf{U}$ とする．定義から，$\mathbf{U}\mathbf{X}=\mathbf{X}$ であって，$\mathbf{U}$ は対称かつベキ等行列（$\mathbf{U}^2=\mathbf{U}$）である．ベクトル $\mathbf{L}$ は $E(\mathbf{L}'\mathbf{Y})=\mathbf{P}'\boldsymbol{\beta}$，つまり $\mathbf{L}'\mathbf{X}=\mathbf{P}'$ であるとする．$\mathbf{U}\mathbf{L}$ と線形関数 $(\mathbf{U}\mathbf{L})'\mathbf{Y}=\mathbf{L}'\mathbf{U}\mathbf{Y}$ を考える．$\mathrm{cov}[\mathbf{L}'(\mathbf{I}-\mathbf{U})\mathbf{Y},\mathbf{L}'\mathbf{U}\mathbf{Y}]=\sigma^2\mathbf{L}'(\mathbf{I}-\mathbf{U})\mathbf{U}\mathbf{L}=0$ であるから

$$E(\mathbf{L}'\mathbf{U}\mathbf{Y})=\mathbf{L}'\mathbf{U}\mathbf{X}\boldsymbol{\beta}=\mathbf{L}'\mathbf{X}\boldsymbol{\beta}=\mathbf{P}'\boldsymbol{\beta}$$
$$V(\mathbf{L}'\mathbf{Y})=V[\mathbf{L}'(\mathbf{I}-\mathbf{U})\mathbf{Y}+\mathbf{L}'\mathbf{U}\mathbf{Y}]$$
$$=V[\mathbf{L}'(\mathbf{I}-\mathbf{U})\mathbf{Y}]+V(\mathbf{L}'\mathbf{U}\mathbf{Y})\geqq V(\mathbf{L}'\mathbf{U}\mathbf{Y}).$$

したがって $\mathbf{L}'\mathbf{U}\mathbf{Y}$ は $\mathbf{P}'\boldsymbol{\beta}$ の不偏推定量であって，$\mathbf{L}'\mathbf{Y}$ よりも小さい分散をもつ．$\mathbf{M}'\mathbf{Y}$ を $\mathbf{P}'\boldsymbol{\beta}$ の他の不偏推定量とすると $(\mathbf{L}'-\mathbf{M}')\mathbf{X}=\mathbf{0}$ であるから，$\mathbf{L}-\mathbf{M}$ は $\mathbf{X}$ の列ベクトルと直交する．したがって $\mathbf{U}(\mathbf{L}-\mathbf{M})=\mathbf{0}$，つまり $\mathbf{M}'\mathbf{U}\mathbf{Y}=\mathbf{L}'\mathbf{U}\mathbf{Y}$ である．それゆえ $V(\mathbf{M}'\mathbf{Y})\geqq V(\mathbf{M}'\mathbf{U}\mathbf{Y})=V(\mathbf{L}'\mathbf{U}\mathbf{Y})$，すなわち $\mathbf{L}'\mathbf{U}\mathbf{Y}$ は l.s. 推定量である．

$\mathbf{X}$ の列空間の上への $\mathbf{L}$ の射影を $\mathbf{X}\boldsymbol{\lambda}$ とすると $\mathbf{X}'(\mathbf{L}-\mathbf{X}\boldsymbol{\lambda})=\mathbf{0}$，すなわち $\boldsymbol{\lambda}$ は方程式 $\mathbf{X}'\mathbf{X}\boldsymbol{\lambda}=\mathbf{X}'\mathbf{L}$ を満足する．したがって $\boldsymbol{\lambda}=(\mathbf{X}'\mathbf{X})^-\mathbf{X}'\mathbf{L}$，$\mathbf{X}\boldsymbol{\lambda}=\mathbf{X}(\mathbf{X}'\mathbf{X})^-\mathbf{X}'\mathbf{L}$ である．ここで $(\mathbf{X}'\mathbf{X})^-$ は $\mathbf{X}'\mathbf{X}$ の一般逆行列である．したがって射影行列 $\mathbf{U}$ を $\mathbf{X}(\mathbf{X}'\mathbf{X})^-\mathbf{X}'$ に選ぶことができる．さて

$$\mathbf{U}\mathbf{L}=\mathbf{X}(\mathbf{X}'\mathbf{X})^-\mathbf{X}'\mathbf{L}=\mathbf{X}(\mathbf{X}'\mathbf{X})^-\mathbf{P}\qquad(\mathbf{P}=\mathbf{X}'\mathbf{L}\ \text{を用いた}),$$
$$\mathbf{L}'\mathbf{U}\mathbf{Y}=\mathbf{P}'(\mathbf{X}'\mathbf{X})^-\mathbf{X}'\mathbf{Y}=\mathbf{P}'\hat{\boldsymbol{\beta}}.$$

かくして $\mathbf{P}'(\mathbf{X}'\mathbf{X})^-\mathbf{X}'\mathbf{Y}$ または $\mathbf{P}'\hat{\boldsymbol{\beta}}$ は $\mathbf{P}'\boldsymbol{\beta}$ の最小2乗推定量である．$\mathbf{P}'[(\mathbf{X}'\mathbf{X})^-]'\mathbf{X}'=\mathbf{P}'(\mathbf{X}'\mathbf{X})^-\mathbf{X}'$ であることを注意しておく．

2次形式の極値　l.s. 推定量の問題は，与えられた母数の関数 $\mathbf{P}'\boldsymbol{\beta}$ にたいして不偏でありかつ最小分散をもつ線形関数 $\mathbf{L}'\mathbf{Y}$ を決定する問題とみなされる．$V(\mathbf{L}'\mathbf{Y})=\sigma^2\mathbf{L}'\mathbf{L}$ であり，$E(\mathbf{L}'\mathbf{Y})=\mathbf{P}'\boldsymbol{\beta}$ から $\mathbf{X}'\mathbf{L}=\mathbf{P}$ が得られるから，問題は，条件 $\mathbf{X}'\mathbf{L}=\mathbf{P}$ のもとで $\mathbf{L}'\mathbf{L}$ を最小にする $\mathbf{L}$ を求めること，または $\mathbf{X}'\mathbf{L}=\mathbf{P}$ の解 $\mathbf{L}$ のうちで最小の長さをも

つものを決定することである．この問題は［(ii), 1f. 1］で考えたものの特別な場合（$\mathbf{A}=\mathbf{I}$ の場合）である．そこでの結果を用いると，$\mathbf{L}'\mathbf{L}$ は

$$(\mathbf{L}_*)' = \mathbf{P}'(\mathbf{X}'\mathbf{X})^-\mathbf{X}'$$

のとき最小となるから，最小分散推定量は

$$(\mathbf{L}_*)'\mathbf{Y} = \mathbf{P}'(\mathbf{X}'\mathbf{X})^-\mathbf{X}'\mathbf{Y} = \mathbf{P}'\hat{\boldsymbol{\beta}}$$

となり，これは l. s. 推定量である．

4a. 7　観測値に相関のある場合の具体的な表現

問題 $(\mathbf{Y}, \mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{G})$ と $(\mathbf{Y}, \mathbf{X}\boldsymbol{\beta}, \boldsymbol{\Sigma})$ は，変量 $\mathbf{Y}$ に適当な変換をほどこすことにより，$(\mathbf{Y}, \mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I})$ に帰着させることができることを 4a. 1 に示した．ここでは，もとの相関のある観測値を用いて正規方程式や l. s. 推定量などの具体的な表現を与える．$(\mathbf{Y}, \mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I})$ の場合の表現における $\mathbf{Y}, \mathbf{X}$ のところに変換した値を代入すると，もとの観測値を用いた具体的な表現が得られる．表 4a. 7 はこの結果をまとめたものである．

本質的なちがいは正規方程式 $\mathbf{S}\boldsymbol{\beta} = \mathbf{Q}$ の計算だけであることがわかる．そのあとの公式は $\mathbf{S}$ と $\mathbf{Q}$ だけに関係している．

4a. 8　最小2乗法の理論における計算上の問題

最小2乗法の適用に際しては，正規方程式における $m \times m$ 行列 $\mathbf{X}'\mathbf{X}$ を数値的に簡約する必要があり，それはある場合には困難をともなう．机上計算および高速計算機の両方の計算において便利である1つの方法は次のものである．まず，$\mathbf{X}'\mathbf{X} = \mathbf{T}'\mathbf{T}$ であるような行列 $\mathbf{T}$ をみつける．ただし $\mathbf{T}$ は m 次の上三角行列である．$\mathbf{X}'\mathbf{X}$ が特異であるときには，$\mathbf{T}$ は，1つの対角要素がゼロならばその行の要素はすべてゼロであるという性質をもつ．ゼロでない対角要素の数は $R(\mathbf{X})$ に等しく，$\mathbf{X}'\mathbf{X}$ が特異であるならばそれは m よりも小さい．このような行列が与えられると，それの g 逆行列 $\mathbf{T}^-$ をみつけることは容易である．これは，まずゼロの行を除き，次にこの簡約された三角行列の真の逆行列を求め，最後に適当なところにゼロの行を挿入して行列の大きさをもとのものと同じにすることによって得られる．あとの計算は行列のかけ算だけである．

$\mathbf{X}'\mathbf{X}$ の1つの一般逆行列は $\mathbf{C} = \mathbf{T}^-(\mathbf{T}^-)'$ であり，正規方程式の解は $\hat{\boldsymbol{\beta}} = \mathbf{C}\mathbf{Q}$ である．推定可能な母数の関数の分散共分散を計算するための $\hat{\boldsymbol{\beta}}$ の分散共分散行列は $\sigma^2\mathbf{C}$ である．σ^2 を推定するために用いられる2乗和の最小値は $R_0^2 = \mathbf{Y}'\mathbf{Y} - \hat{\boldsymbol{\beta}}'\mathbf{Q}$ である．

$\mathbf{T}$ の計算には数値的に安定した方法を用いるべきである．行列 $\mathbf{X}'\mathbf{X}$ が特に悪い条件でないならば，普通には平方根法がよいであろう．よりよい他の方法は，$\mathbf{X}$ に左からハウスホルダーの直交行列をかけて三角行列にもっていく方法である［(vii), 1b. 2 を参照］．

表 4a. 7

問題	$(\mathbf{Y}, \mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I})$	$(\mathbf{Y}, \mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{G})$	$(\mathbf{Y}, \mathbf{X}\boldsymbol{\beta}, \boldsymbol{\Sigma})$
変換	$\mathbf{Y} \to \mathbf{Y}$	$\mathbf{Y} \to \mathbf{G}^{-\frac{1}{2}}\mathbf{Y},\ \mathbf{X} \to \mathbf{G}^{-\frac{1}{2}}\mathbf{X}$	$\mathbf{Y} \to \boldsymbol{\Sigma}^{-\frac{1}{2}}\mathbf{Y},\ \mathbf{X} \to \boldsymbol{\Sigma}^{-\frac{1}{2}}\mathbf{X}$
最小にするべき2次形式	$(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})$	$(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})'\mathbf{G}^{-1}(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})$	$(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})'\boldsymbol{\Sigma}^{-1}(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})$
正規方程式 $(\mathbf{S}\boldsymbol{\beta} = \mathbf{Q})$	$\mathbf{S} = \mathbf{X}'\mathbf{X},\ \mathbf{Q} = \mathbf{X}'\mathbf{Y}$	$\mathbf{S} = \mathbf{X}'\mathbf{G}^{-1}\mathbf{X},\ \mathbf{Q} = \mathbf{X}'\mathbf{G}^{-1}\mathbf{Y}$	$\mathbf{S} = \mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X},\ \mathbf{Q} = \mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{Y}$
解 $\hat{\boldsymbol{\beta}}$ $(\mathbf{S}^- = \mathbf{C})$	$\hat{\boldsymbol{\beta}} = \mathbf{C}\mathbf{Q}$	$\hat{\boldsymbol{\beta}} = \mathbf{C}\mathbf{Q}$	$\hat{\boldsymbol{\beta}} = \mathbf{C}\mathbf{Q}$
$\hat{\boldsymbol{\beta}}$ の形式的な分散共分散行列	$\sigma^2\mathbf{C}$	$\sigma^2\mathbf{C}$	$\mathbf{C}$
残差平方和 R_0^2	$\mathbf{Y}'\mathbf{Y} - \mathbf{Q}'\hat{\boldsymbol{\beta}}$	$\mathbf{Y}'\mathbf{G}^{-1}\mathbf{Y} - \mathbf{Q}'\hat{\boldsymbol{\beta}}$	$\mathbf{Y}'\boldsymbol{\Sigma}^{-1}\mathbf{Y} - \mathbf{Q}'\hat{\boldsymbol{\beta}}$
$E(R_0^2)$	$(n-r)\sigma^2$	$(n-r)\sigma^2$	$(n-r)$
$\mathbf{P}'\boldsymbol{\beta}$ が推定可能であるための条件	$\mathbf{X}'\mathbf{L} = \mathbf{P}$	$\mathbf{X}'\mathbf{L} = \mathbf{P}$	$\mathbf{X}'\mathbf{L} = \mathbf{P}$
$\mathbf{P}'\hat{\boldsymbol{\beta}}$ の分散	$\sigma^2\mathbf{P}'\mathbf{C}\mathbf{P}$	$\sigma^2\mathbf{P}'\mathbf{C}\mathbf{P}$	$\mathbf{P}'\mathbf{C}\mathbf{P}$
$\mathbf{P}'\boldsymbol{\beta}$, $\mathbf{R}'\hat{\boldsymbol{\beta}}$ の共分散	$\sigma^2\mathbf{P}'\mathbf{C}\mathbf{R}$	$\sigma^2\mathbf{P}'\mathbf{C}\mathbf{R}$	$\mathbf{P}'\mathbf{C}\mathbf{R}$

（$\mathbf{G}$ と $\boldsymbol{\Sigma}$ は正則であると仮定する）

4a.9　母数に制約があるときの最小2乗推定

問題 $(\mathbf{Y}, \mathbf{X}\beta, \sigma^2\mathbf{I})$ において，β に解をもつ線形の制約式 $\mathbf{H}'\beta = \xi$（既知）がある場合を考えよう．ただし $\operatorname{rank}\mathbf{H} = m - s$ とする．$\mathbf{H}'\beta = \xi$ の一般解は $\beta_0 + \mathbf{B}\theta$ である．ただし β_0 は特解，$\mathbf{B}$ は $\mathbf{H}'\mathbf{B} = 0$ であるような階数が s の $m \times s$ 行列，θ は s 次元の任意のベクトルで，これがいわば新しい母数である．

$$E(\mathbf{Y}) = \mathbf{X}\beta = \mathbf{X}(\beta_0 + \mathbf{B}\theta) \quad \text{または} \quad E(\mathbf{Y} - \mathbf{X}\beta_0) = \mathbf{X}\mathbf{B}\theta.$$

$\mathbf{Z} = \mathbf{Y} - \mathbf{X}\beta_0$ とおくと $E(\mathbf{Z}) = \mathbf{X}\mathbf{B}\theta$ であり，新しい変量 $\mathbf{Z}$ と新しい母数 θ を用いると問題は $(\mathbf{Z}, \mathbf{X}\mathbf{B}\theta, \sigma^2\mathbf{I})$ となり，(4a.1.2) の形に帰着された．$\mathbf{P}'\beta = \mathbf{P}'(\beta_0 + \mathbf{B}\theta) = \mathbf{P}'\beta_0 + \mathbf{P}'\mathbf{B}\theta$ を推定するためには，$\mathbf{P}'\mathbf{B}\theta = \mathbf{R}'\theta$ を考える必要がある．4a.5 のように，σ^2 の推定にたいしては $\mathbf{F} = \mathbf{X}\mathbf{B}$ の階数が必要である．$\mathbf{H}'\mathbf{B} = 0$ を用いると

$$\operatorname{rank}(\mathbf{X}' \,\vdots\, \mathbf{H}) = \operatorname{rank}\mathbf{X}\mathbf{B} + \operatorname{rank}\mathbf{H} \tag{4a.9.1}$$

が容易に示される．したがって実際に $\mathbf{B}$ や $\mathbf{F}$ を計算しなくても，行列 $\mathbf{X}$ と $\mathbf{H}$ をしらべることにより (4a.9.1) から $\mathbf{X}\mathbf{B}$ の階数は求められる．

しかし2次形式の極値に関する［(iii)，1f.1］の結果を使うと，減らした母数を導入しなくてこの問題をしらべることができる．

問題は，条件

$$\mathbf{H}'\beta = \xi \quad \text{であって} \quad E(\mathbf{L}'\mathbf{Y} + d) = \mathbf{P}'\beta \tag{4a.9.2}$$

のもとで

$$V(\mathbf{L}'\mathbf{Y} + d) = \sigma^2\mathbf{L}'\mathbf{L} \quad \text{が最小}$$

であるような線形関数 $\mathbf{L}'\mathbf{Y} + d$ をみいだすこととみなしてよい．条件 (4a.9.2) は，方程式

$$\mathbf{L}'\mathbf{X}\beta + d = \mathbf{P}'\beta \tag{4a.9.3}$$

が，$\mathbf{H}'\beta - \xi = 0$ であるようなすべての β にたいして満足されるということを導く．そうすると

$$\mathbf{L}'\mathbf{X} + \mathbf{M}'\mathbf{H}' = \mathbf{P}' \quad \text{または} \quad (\mathbf{X}' \,\vdots\, \mathbf{H})\begin{pmatrix}\mathbf{L}\\ \cdots\\ \mathbf{M}\end{pmatrix} = \mathbf{P} \tag{4a.9.4}$$

であって $d = \mathbf{M}'\xi$ であるようなベクトル $\mathbf{M}$ が存在する．かくして問題は

$$(\mathbf{X}' \,\vdots\, \mathbf{H})\begin{pmatrix}\mathbf{L}\\ \cdots\\ \mathbf{M}\end{pmatrix} = \mathbf{P} \quad \text{という条件のもとで} \quad \mathbf{L}'\mathbf{L} \quad \text{を最小にすること} \tag{4a.9.5}$$

となる．［(iii)，1f.1］の結果を使うと，$\mathbf{L}, \mathbf{M}$ の最適な選択は

$$\mathbf{L}_* = \mathbf{X}\mathbf{C}_1'\mathbf{P}, \qquad \mathbf{M}_* = \mathbf{C}_2'\mathbf{P} \tag{4a.9.6}$$

である．ただし $\mathbf{C}_1, \mathbf{C}_2$ は一般逆行列

$$\begin{pmatrix}\mathbf{X}'\mathbf{X} & \mathbf{H}\\ \mathbf{H}' & 0\end{pmatrix}^- = \begin{pmatrix}\mathbf{C}_1 & \mathbf{C}_2\\ \mathbf{C}_3 & \mathbf{C}_4\end{pmatrix} \tag{4a.9.7}$$

の部分行列である．したがって次の結果を得る．

$\mathbf{P}'\beta$ の最小分散不偏推定量は $(\mathbf{L}_*)'\mathbf{Y} + (\mathbf{M}_*)'\xi = \mathbf{P}'(\mathbf{C}_1\mathbf{X}'\mathbf{Y} + \mathbf{C}_2\xi)$ であり，これは方程式

$$\begin{aligned}\mathbf{X}'\mathbf{X}\beta + \mathbf{H}\lambda &= \mathbf{X}'\mathbf{Y},\\ \mathbf{H}'\beta \qquad &= \xi\end{aligned} \tag{4a.9.8}$$

の解を $\hat{\beta}$ とするとき $\mathbf{P}'\hat{\beta}$ と書いてもよい．

方程式 (4a.9.8) は，ラグランジュ乗数 λ を用い，制約式 $\mathbf{H}'\beta = \xi$ のもとで $(\mathbf{Y} - \mathbf{X}\beta)'(\mathbf{Y} - \mathbf{X}\beta)$ を最小にすることによって得られる正規方程式と同じである．

推定量の分散共分散　$\mathbf{P}'\hat{\beta} = \mathbf{P}'\mathbf{C}_1\mathbf{X}'\mathbf{Y} + \mathbf{P}'\mathbf{C}_2\xi$ であるから，一般逆行列 (4a.9.7) の性質を用いると

$$V(\mathbf{P}'\hat{\beta}) = \sigma^2\mathbf{P}'\mathbf{C}_1\mathbf{X}'\mathbf{X}\mathbf{C}_1\mathbf{P} = \sigma^2\mathbf{P}'\mathbf{C}_1\mathbf{P}$$

である．同様にして

$$\operatorname{cov}(\mathbf{P}'\hat{\beta}, \mathbf{R}'\hat{\beta}) = \sigma^2\mathbf{P}'\mathbf{C}_1\mathbf{R} = \sigma^2\mathbf{R}'\mathbf{C}_1\mathbf{P}.$$

行列 $\mathbf{C}_1$ は計算されないが，$\mathbf{P}'\hat{\beta} = \gamma'\mathbf{X}'\mathbf{Y} + \delta'\xi$，$\mathbf{R}'\hat{\beta} = \phi'\mathbf{X}'\mathbf{Y} + \psi'\xi$，$\mathbf{H}'\gamma = \mathbf{H}'\phi = 0$ ということがわかっているとする．そうすると分散共分散

$$V(\mathbf{P}'\hat{\beta}) = \sigma^2\mathbf{P}'\gamma, \qquad V(\mathbf{R}'\hat{\beta}) = \sigma^2\mathbf{R}'\phi$$
$$\operatorname{cov}(\mathbf{P}'\hat{\beta}, \mathbf{R}'\hat{\beta}) = \sigma^2\mathbf{P}'\phi = \sigma^2\mathbf{R}'\gamma$$

はベクトル $\mathbf{P}, \mathbf{R}, \gamma, \phi$ だけを含んでいる．もしも条件 $\mathbf{H}'\gamma = \mathbf{H}'\phi = 0$ が満足されないならば

$$V(\mathbf{P}'\hat{\beta}) = \sigma^2\gamma'\mathbf{X}'\mathbf{X}\gamma, \qquad \operatorname{cov}(\mathbf{P}'\hat{\beta}, \mathbf{R}'\hat{\beta}) = \sigma^2\gamma'\mathbf{X}'\mathbf{X}\phi.$$

σ^2 の推定　σ^2 の不偏推定量は

$$(\mathbf{Y}'\mathbf{Y} - \hat{\beta}'\mathbf{X}'\mathbf{Y} - \hat{\lambda}'\xi) \div (n - s)$$

である．ただし，(4a.9.1) で示したように $s = \operatorname{rank}(\mathbf{X}' \,\vdots\, \mathbf{H}) - \operatorname{rank}\mathbf{H}$ であり，$(\hat{\beta}, \hat{\lambda})$ は (4a.9.8) の任意の解である．一般逆行列 (4a.9.7) の部分行列を使うと

$$\hat{\beta} = \mathbf{C}_1\mathbf{X}'\mathbf{Y} + \mathbf{C}_2\xi, \qquad \hat{\lambda} = \mathbf{C}_3\mathbf{X}'\mathbf{Y} + \mathbf{C}_4\xi$$

であることがわかる．

4a.10　母数の関数の同時推定

母数 β に制約のある場合やない場合を含め一般的な模型 $(\mathbf{Y}, \mathbf{X}\beta, \Sigma)$ を考えよう．$\mathbf{P}_1'\hat{\beta}, \cdots, \mathbf{P}_k'\hat{\beta}$ をそれぞれ母数の関数 $\mathbf{P}_1'\beta, \cdots, \mathbf{P}_k'\beta$ の最小2乗推定量とする．さらに，$\mathbf{A}$ を推定量 $\mathbf{P}_1'\hat{\beta}, \cdots, \mathbf{P}_k'\hat{\beta}$ の分散共分散行列とする．そうすると，l.s. 推定量について次の最適性が得られる．

(i)　$\mathbf{L}_1'\mathbf{Y}, \cdots, \mathbf{L}_k'\mathbf{Y}$ を $\mathbf{P}_1'\beta, \cdots, \mathbf{P}_k'\beta$ の任意の不偏推定量とし，それの分散共分散行列

を $\mathbf{B}$ とする．そうすると $\mathbf{B}-\mathbf{A}$ は非負定符号であり，したがって

（a）$\mathrm{Trace}\,\mathbf{B} \geqslant \mathrm{Trace}\,\mathbf{A}$

（b）$|\mathbf{B}| \geqslant |\mathbf{A}|$

（c）$\mathbf{Q}$ を任意の非負定符号行列とすると，$\mathrm{Trace}\,\mathbf{QB} \geqslant \mathrm{Trace}\,\mathbf{QA}$

（d）$\mathbf{B}$ の最大固有値 $\geqslant$ $\mathbf{A}$ の最大固有値

である．

母数の線形関数

$$a_1\mathbf{P}_1'\beta + \cdots + a_k\mathbf{P}_k'\beta$$

を考えよう．それの l. s. 推定量は $a_1\mathbf{P}_1'\hat{\beta} + \cdots + a_k\mathbf{P}_k'\hat{\beta}$ であり，その分散は $\mathbf{a}'\mathbf{Aa}$ である．ただし，$\mathbf{a}' = (a_1, \cdots, a_k)$．他の推定量 $a_1\mathbf{L}_1'\mathbf{Y} + \cdots + a_k\mathbf{L}_k'\mathbf{Y}$ の分散は $\mathbf{a}'\mathbf{Ba}$ である．したがって，すべての $\mathbf{a}$ にたいして $\mathbf{a}'\mathbf{Ba} \geqq \mathbf{a}'\mathbf{Aa}$ が得られ，これから (a)—(d) の結果が証明される．

4a. 11　母数が確率変数であるときの最小2乗法の理論

$(\mathbf{Y}, \mathbf{X}\beta, \Sigma)$ において，β 自身が平均 μ，分散共分散行列 $\mathbf{T}$ をもつ確率変数である場合を考えよう．確率変数 $\mathbf{P}'\beta$ の最良線形推定量（または予測量）は何であるか．$\Sigma = D(\mathbf{Y}|\beta)$，つまり Σ は β が与えられたときの $\mathbf{Y}$ の条件付分散共分散行列であることに注意をせよ．これより全分散共分散行列，全共分散行列と条件付のそれらの行列との関係をあらわす次の公式が得られる．

$$\begin{aligned} D(\mathbf{Y}) &= ED(\mathbf{Y}|\beta) + D[E(\mathbf{Y}|\beta)] \\ &= \Sigma + D(\mathbf{X}\beta) = \Sigma + \mathbf{XTX}' \end{aligned} \tag{4a. 11. 1}$$

$$\begin{aligned} C(\mathbf{Y}, \mathbf{P}'\beta) &= E[C(\mathbf{Y}, \mathbf{P}'\beta)|\beta] + C[E(\mathbf{Y}|\beta), \mathbf{P}'\beta] \\ &= 0 + \mathbf{XTP}. \end{aligned} \tag{4a. 11. 2}$$

さて

$$\left.\begin{aligned} &E(\mathbf{P}'\beta - a - \mathbf{L}'\mathbf{Y}) = 0 \\ &V(\mathbf{P}'\beta - a - \mathbf{L}'\mathbf{Y}) \text{ が最小} \end{aligned}\right\} \tag{4a. 11. 3}$$

であるような線形関数 $a + \mathbf{L}'\mathbf{Y}$ を決定しよう．

（i）場合1，μ 既知　$\mathbf{L}, a$ の最適な選択は

$$\begin{aligned} \mathbf{L}_* &= (\Sigma + \mathbf{XTX}')^{-1}\mathbf{XTP} = \Sigma^{-1}\mathbf{X}(\mathbf{T}^{-1} + \mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{P} \\ a_* &= \mu'\mathbf{P} - \mu'\mathbf{X}'\mathbf{L}_* \end{aligned} \tag{4a. 11. 4}$$

であり，それの予測分散は

$$\begin{aligned} V(\mathbf{P}'\beta - a_* - (\mathbf{L}_*)'\mathbf{Y}) &= \mathbf{P}'\mathbf{TP} - \mathbf{P}'\mathbf{TX}'\mathbf{L}_* \\ &= \mathbf{P}'(\mathbf{T}^{-1} + \mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{P} \end{aligned} \tag{4a. 11. 5}$$

である．

証明は，選択 (4a. 11. 4) が (4a. 11. 3) の最初の条件を満足し，$\mathbf{L}, a$ の任意の他の選択にたいして

$$\begin{aligned} V(\mathbf{P}'\beta - a - \mathbf{L}'\mathbf{Y}) &= V(\mathbf{P}'\beta - a_* - (\mathbf{L}_*)'\mathbf{Y}) + V[(\mathbf{L} - \mathbf{L}_*)'\mathbf{Y}] \\ &\geqslant V(\mathbf{P}'\beta - a_* - (\mathbf{L}_*)'\mathbf{Y}) \end{aligned} \tag{4a. 11. 6}$$

であることを確かめることにある．

（ii）場合2，μ 未知　$\mathbf{X}'\mathbf{L} = \mathbf{P}$ であるような $\mathbf{L}$ が存在するならば，$\mathbf{L}, a$ の最適な選択は

$$\mathbf{L}_* = \Sigma^{-1}\mathbf{X}(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{P}, \qquad a_* = 0 \tag{4a. 11. 7}$$

であり，それの予測分散は

$$\mathbf{P}'(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{P} \tag{4a. 11. 8}$$

である．

最良推定量と予測分散は，(4a. 1. 2) におけるように，β を母数と考えた場合の l. s. 推定量とそれの分散と同じであるということに注意をせよ．

(4a. 11. 8) においては $(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}$ が存在するということを仮定している．そうでないときには $\mathbf{X}'\Sigma^{-1}\mathbf{X}$ の g 逆行列を代入すればよい．

$\mathbf{X}'\mathbf{L}_* = \mathbf{P}$ であることが容易に確かめられるから，条件 (4a. 11. 3) を満足する．等式

$$V[\mathbf{P}'\beta - (\mathbf{L}_*)'\mathbf{Y} - (\mathbf{L} - \mathbf{L}_*)'\mathbf{Y}] = V[\mathbf{P}'\beta - (\mathbf{L}_*)'\mathbf{Y}] + V[(\mathbf{L} - \mathbf{L}_*)'\mathbf{Y}] \tag{4a. 11. 9}$$

が成立する．それは，条件 $\mathbf{L}'\mathbf{X} = (\mathbf{L}_*)'\mathbf{X} = \mathbf{P}$ を用いると，共分散の項は下に示すようにゼロになるからである．

$$\begin{aligned} C[(\mathbf{P}'\beta - (\mathbf{L}_*)'\mathbf{Y}), (\mathbf{L} - \mathbf{L}_*)'\mathbf{Y}] &= (\mathbf{L} - \mathbf{L}_*)'\mathbf{XTP} - (\mathbf{L} - \mathbf{L}_*)'\Sigma\mathbf{L}_* \\ &= 0 - (\mathbf{L} - \mathbf{L}_*)'\Sigma\Sigma^{-1}\mathbf{X}(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{P} = 0 - 0. \end{aligned} \tag{4a. 11. 10}$$

等式 (4a. 11. 9) は求める結果を示している．予測分散が (4a. 11. 8) に与えたものになることは容易に証明される．

4a. 12　計画行列の選択

これまで，$\mathbf{X}$ の選択は実際にはどのようにしてなされるかということを議論しなかった．$\mathbf{X}$ の各行は，ある応答（実験の結果）が観測されたところの“因子”（または，その値は任意に指定することができて実験の結果ではないような器械類の変数）の組合せをあらわす．行列 $\mathbf{X}$ は前もってきめられ，計画行列と呼ばれる．模型 $(\mathbf{Y}, \mathbf{X}\beta, \Sigma)$ では，期待応答は因子変数の線形関数であることが仮定されている．

与えられた $\mathbf{X}$ にたいして，l. s. 推定量 $\hat{\beta}_1, \cdots, \hat{\beta}_m$ の分散共分散行列は $(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}$（ま

たは最大階数でないときには g 逆行列）であることが示されている．さて $(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}$ をできるだけ小さくするためには $\mathbf{X}$ をどのように選ぶべきであるかという問題をとりあげよう．

もし因子の値をまったく自由な範囲で選ぶことが許されるならば，$(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}$ の要素は任意に小さくすることができる．なぜならば，$\mathbf{X}$ を $a\mathbf{X}$ に選ぶことにより，$(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}$ は $a^{-2}(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}$ に変わり，それは $a \to \infty$ のとき $\to 0$ であるからである．しかし実際にはそうすることは誤りであろう．というのは，応答模型は因子の値の広い範囲にたいしては正しくないであろうからである．その範囲は，われわれに興味があり，また応答関数を探究したい範囲内の小さい領域に制限をしなければいけない．

したがって

$$\mathbf{X}_i'\Sigma^{-1}\mathbf{X}_i = C_i^2 \quad (\text{ある定められた値}), \qquad i = 1, \cdots, m \tag{4a. 12. 1}$$

という形の制約をおく．ここで $\mathbf{X}_i$ は $\mathbf{X}$ の列ベクトルであり，実験に用いられた i 番目の因子の n 個の水準をあらわす．条件 (4a. 12. 1) のもとで，可能な最小の分散をもつ $\beta_1, \cdots, \beta_m$ の推定量が得られるように因子の値の組合せまたは $\mathbf{X}$ の行を選びたい．

（i）$\mathbf{X}$ を計画行列とし，$\hat{\beta}_i$ を β_i の l. s. 推定量とする．そうすると，$\mathbf{X}$ についての条件 (4a. 12. 1) のもとで

$$V(\hat{\beta}_i) \geqslant \frac{1}{C_i^2} \tag{4a. 12. 2}$$

であり，最小値は $\mathbf{X}_i'\Sigma^{-1}\mathbf{X}_j = 0,\ j = 1, \cdots, i-1, i+1, \cdots, m$ のとき到達される．

$i = 1$ にたいしてこの結果を証明しよう．行列 $\mathbf{X}'\Sigma^{-1}\mathbf{X}$ は分割形にして

$$\mathbf{X}'\Sigma^{-1}\mathbf{X} = \begin{pmatrix} \mathbf{X}_1'\Sigma^{-1}\mathbf{X}_1 & \mathbf{B}' \\ \mathbf{B} & \mathbf{F} \end{pmatrix} \tag{4a. 12. 3}$$

と書くことができる．したがって $|\mathbf{X}'\Sigma^{-1}\mathbf{X}| = |\mathbf{F}|(\mathbf{X}_1'\Sigma^{-1}\mathbf{X}_1 - \mathbf{B}'\mathbf{F}^{-1}\mathbf{B})$ であり，$\mathbf{B}'\mathbf{F}^{-1}\mathbf{B}$ が非負であることから

$$\frac{|\mathbf{F}|}{|\mathbf{X}'\Sigma^{-1}\mathbf{X}|} = \frac{1}{\mathbf{X}_1'\Sigma^{-1}\mathbf{X}_1 - \mathbf{B}'\mathbf{F}^{-1}\mathbf{B}} \geqslant \frac{1}{\mathbf{X}_1'\Sigma^{-1}\mathbf{X}_1} \tag{4a. 12. 4}$$

が得られる．しかし $V(\hat{\beta}_1) = |\mathbf{F}|/|\mathbf{X}'\Sigma^{-1}\mathbf{X}|$ である．したがって $V(\hat{\beta}_1) \geqslant 1/C_1^2$．(4a. 12. 4) における等式は $\mathbf{B} = 0$ のとき，つまり $\mathbf{X}_1'\Sigma^{-1}\mathbf{X}_j = 0,\ j \neq 1$ のとき到達される．$\Sigma = \sigma^2\mathbf{I}$ の特別な場合には次の結果が得られる．

（ii）条件 $\mathbf{X}_i'\mathbf{X}_i = C_i^2$ のもとでは，組合せ（$\mathbf{X}$ の行）の最適な選択は $\mathbf{X}_i'\mathbf{X}_j = 0$ のとき，つまり $\mathbf{X}$ の列が直交するときである．

4b　仮説検定と区間推定

母数の関数の l. s. 推定値は単に点推定値であって，それと母数の関数の真値との偏差の確率についての陳述は，考えている変量 $\mathbf{Y}$ の分布を規定しなくてはすることはできない．$\mathbf{Y} \sim N(\mathbf{X}\beta, \sigma^2\mathbf{I})$ を仮定する．この模型のもとでのいろいろな統計量の標本分布は，3b. 5 で調べた．$\mathbf{H}'$ を階数 k の大きさ $k \times m$ の行列とし，その各行は $\mathbf{X}$ の行に従属しているとする．このことは，k 個の母数の関数 $\mathbf{H}'\beta$ のおのおのは模型 $(\mathbf{Y}, \mathbf{X}\beta, \sigma^2\mathbf{I})$ のもとで推定可能であり，そしてこの場合には l. s. 推定量は $\mathbf{H}'\hat{\beta} = \mathbf{Z}$（とおく）であるということを意味する [(i), 4a. 2 を参照]．$\mathbf{X}, \mathbf{Y}$ であらわすと，$\mathbf{H}'\hat{\beta} = \mathbf{H}'(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'\mathbf{Y}$，$R_0^2 = \mathbf{Y}'(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}')\mathbf{Y}$ である．そうすると

$$\mathbf{H}'(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}') = 0$$

であり，このことは $\mathbf{H}'\hat{\beta}$ と R_0^2 とが独立に分布することを示している（第3章の終りの問題 2. 2, 2. 3 をも参照せよ）．

(3b. 3. 4) から $\mathbf{Z}$ の分布は平均 $\mathbf{H}'\beta$，分散共分散行列 $\sigma^2\mathbf{D}$（とおく）の k 次元正規であり，R_0^2 の分布は $\sigma^2\chi^2(n-r)$ である．この結果を母数の関数の未知な値の推測に用いよう．

4b. 1　母数の1つの関数（の推測）

推定可能な母数の関数 $\theta = \mathbf{P}'\beta$ が与えられ，それの l. s. 推定量を u とし，その分散を $p^2\sigma^2$ とおく．そうすると

$$u \sim N(\theta, p^2\sigma^2) \quad と \quad R_0^2 \sim \sigma^2\chi^2(n-r)$$

は独立であるから，$s^2 = R_0^2/(n-r)$ とおくと

$$t = \frac{u-\theta}{p\sigma} \div \frac{s}{\sigma} = \frac{u-\theta}{ps} \sim S(n-r), \tag{4b. 1. 1}$$

つまり t は自由度 $n-r$ のスチューデントの分布に従う [(3b. 1. 8) を用いた]．t_α を $|t|$ の α 確率点，つまり $P(|t| > t_\alpha) = \alpha$ とすると，

$$P\left\{\frac{|u-\theta|}{ps} \leqslant t_\alpha\right\} = 1 - \alpha, \tag{4b. 1. 2}$$

すなわち

$$P(u - pst_\alpha \leqslant \theta \leqslant u + pst_\alpha) = 1 - \alpha \tag{4b. 1. 3}$$

である．結果 (4b. 1. 2) と (4b. 1. 3) は，与えられた母数の関数の値 θ についての推測をするのに非常に重要である．

仮説検定　帰無仮説を $\mathbf{P}'\beta = \theta_0$（1つの与えられた値）とする．帰無仮説が真であるならば，(4b. 1. 2) を用いると，

$$P\left\{\frac{|u-\theta_0|}{ps} \leqslant t_\alpha\right\} = 1-\alpha \tag{4b. 1. 4}$$

が得られる．かくして有意水準 α では，観測された u にたいして

$$\frac{|u-\theta_0|}{ps} > t_\alpha \tag{4b. 1. 5}$$

のとき帰無仮説は棄却される．

区間推定　(4b. 1. 3) は，観測された u にたいして，区間

$$(u - pst_\alpha, u + pst_\alpha) \tag{4b. 1. 6}$$

が θ の $1-\alpha$ 信頼区間であることを導く（信頼区間のくわしい取り扱いについては 7b を参照）．

4b. 2　母数の2つ以上の関数（の推測）

k 個の独立かつ推定可能な母数の関数

$$\theta_1 = \mathbf{H}_1'\beta, \quad \cdots, \quad \theta_k = \mathbf{H}_k'\beta \tag{4b. 2. 1}$$

を考えよう．それらは行列記号では，$\theta = \mathbf{H}'\beta$ と書ける．ただし $\mathbf{H}'$ は $k \times m$ の行列であり，θ は $\theta_1, \cdots, \theta_k$ の列ベクトルである．$(\theta_1, \cdots, \theta_k)$ の l. s. 推定量は $(z_1, \cdots, z_k) = \mathbf{Z}$ であらわされ，それの分散共分散行列は $\sigma^2\mathbf{D}$ であらわされる．そうすると $E(\mathbf{Z}) = \theta$, $D(\mathbf{Z}) = \sigma^2\mathbf{D}$ であるから，(3b. 4. 7) を使うと，

$$(\mathbf{Z}-\theta)'\mathbf{D}^{-1}(\mathbf{Z}-\theta) \sim \sigma^2\chi^2(k), \quad R_0^2 \sim \sigma^2\chi^2(n-r)$$

であり，またそれらは独立であることがわかる．したがって (3b. 1. 11) より

$$F = \frac{(\mathbf{Z}-\theta)'\mathbf{D}^{-1}(\mathbf{Z}-\theta)}{k} \div \frac{R_0^2}{n-r} \sim F(k, n-r). \tag{4b. 2. 2}$$

統計量 F とその分布 (4b. 2. 2) は，いくつかの母数の関数についての同時推測において基本的な役割を果たす．この統計量について考察をしよう．

多重仮説の検定　k 個の母数の関数がある与えられた値をもつ，すなわち

$$\mathbf{H}_1'\beta = \theta_{10}, \quad \cdots, \quad \mathbf{H}_k'\beta = \theta_{k0}$$

という帰無仮説を検定したいとしよう．この仮説は行列記号では $\mathbf{H}'\beta = \theta_0$ と書ける．l. s. 推定量と与えられた値との偏差のベクトルは $\mathbf{Z} - \theta_0$ である．事実，θ_0 が真でないならば偏差 $\mathbf{Z} - \theta_0$ は大きくなるであろう．偏差を複合したもの（偏差の1次元測度）

$$(\mathbf{Z}-\theta_0)'\mathbf{D}^{-1}(\mathbf{Z}-\theta_0)$$

を考えよう．公式 (4a. 1. 7) を用いると

$$\frac{E(\mathbf{Z}-\theta_0)'\mathbf{D}^{-1}(\mathbf{Z}-\theta_0)}{k}$$

$$= \frac{\sigma^2}{k}\mathrm{trace}(\mathbf{D}^{-1}\mathbf{D}) + \frac{1}{k}E(\mathbf{Z}-\theta_0)\mathbf{D}^{-1}E(\mathbf{Z}-\theta_0)$$

$$= \sigma^2 + \frac{1}{k}(\mathbf{H}'\beta-\theta_0)\mathbf{D}^{-1}(\mathbf{H}'\beta-\theta_0)$$

$$= \sigma^2, \quad \mathbf{H}'\beta = \theta_0 \text{ のとき，すなわち仮説が真のときに限り} \tag{4b. 2. 3}$$

である．

$(\mathbf{H}'\beta-\theta_0)\mathbf{D}^{-1}(\mathbf{H}'\beta-\theta_0)$ は正値であるから，(4b. 2. 3) が σ^2 をこえるときは仮説が正しくないときだけである．一方，$s^2 = [R_0^2/(n-r)]$ の期待値は，未知母数 β について何の仮定もなくて σ^2 である．したがって比

$$F = \frac{(\mathbf{Z}-\theta_0)'\mathbf{D}^{-1}(\mathbf{Z}-\theta_0)}{k} \div \frac{R_0^2}{n-r} \tag{4b. 2. 4}$$

において，分子は，帰無仮説が真でないならば，帰無仮説のちがいをあらわす量 $(\mathbf{H}'\beta-\theta_0)\mathbf{D}^{-1}(\mathbf{H}'\beta-\theta_0)/k$ に関係して，分母よりも平均的に大きい値をもつ．かくして F の大きい値は帰無仮説からのズレを示す．

$$P(F > F_\alpha) = \alpha, \qquad \text{ただし} \quad F \sim F(k, n-r)$$

であるように値 F_α を選ぶと，有意水準 α で仮説を棄却するには，検定基準を

$$F = \frac{(\mathbf{Z}-\theta_0)'\mathbf{D}^{-1}(\mathbf{Z}-\theta_0)}{ks^2} > F_\alpha \tag{4b. 2. 5}$$

とすればよい．

F の計算に際しては，個々の母数の関数の l. s. 推定値を求めることなく，直接分子を与える他の方法がある．

$$R_1^2 = \min_{\mathbf{H}'\beta=\theta_0} (\mathbf{Y}-\mathbf{X}\beta)'(\mathbf{Y}-\mathbf{X}\beta)$$

$$= (\mathbf{Z}-\theta_0)'\mathbf{D}^{-1}(\mathbf{Z}-\theta_0) + R_0^2 \tag{4b. 2. 6}$$

を [(ii), 3b. 5] や第3章の問題 2. 4 で示した．もし R_1^2 が条件 $\mathbf{H}'\beta = \theta_0$ のもとでの β についての制限つき最小値として直接得られるならば，(4b. 2. 6) は公式

$$(\mathbf{Z}-\theta_0)'\mathbf{D}^{-1}(\mathbf{Z}-\theta_0) = R_1^2 - R_0^2 \tag{4b. 2. 7}$$

を与える．したがって検定基準 (4b. 2. 4) を

$$F = \frac{R_1^2 - R_0^2}{k} \div \frac{R_0^2}{n-r} \tag{4b. 2. 8}$$

と書いてもよい．

等式 (4b. 2. 7) は，l. s. 推定量と仮説の値との偏差を複合した測度として，(4b. 2. 8) の F 統計量の分子にあらわれている $R_1^2 - R_0^2$ の具体的表現を与えるという意味において重要である．

F 統計量 (4b. 2. 8) を求める計算は，分散分析と呼ばれる表の形で与えられる．＊印のところの数値は引き算で求める．F 統計値は M. S. 値の比である．

同時信頼区間 (Scheffe (1959))　(4b. 2. 2) から

表 4b. 2　分散分析

	D. F.	平方和 (S. S.)	平均平方 (M. S.)
仮説 $\mathbf{H}'\beta=\theta_0$ からの偏差	k	$*=R_1^2-R_0^2$	$\dfrac{R_1^2-R_0^2}{k}$
残　差	$n-r$	$\min\limits_{\beta}(\mathbf{Y}-\mathbf{X}\beta)'(\mathbf{Y}-\mathbf{X}\beta)=R_0^2$	$\dfrac{R_0^2}{n-r}$
合　計	$n-r+k$	$\min\limits_{\mathbf{H}'\beta=\theta_0}(\mathbf{Y}-\mathbf{X}\beta)'(\mathbf{Y}-\mathbf{X}\beta)=R_1^2$	

$$P\left\{\frac{(\mathbf{Z}-\theta)'\mathbf{D}^{-1}(\mathbf{Z}-\theta)}{ks^2}\leqslant F_\alpha\right\}=1-\alpha \tag{4b. 2. 9}$$

が得られる．ただし $\theta=\mathbf{H}'\beta$ は β の k 個の独立で推定可能な母数の関数をあらわす．与えられた $\mathbf{Z}, s^2$ にたいして，(4b. 2. 9) の括弧のなかの不等式を満足するすべての可能な θ の集合は，閉領域であり，それは $\theta=\mathbf{H}'\beta$ の $1-\alpha$ 信頼域と呼ばれる．このような集合を C であらわそう．そうすると関数 $g_i(\theta)$, $i=1,2,\cdots,k$ の信頼係数 $1-\alpha$ 以上の同時信頼区間は

$$[\min_{\theta\in C} g_i(\theta), \max_{\theta\in C} g_i(\theta)],\qquad i=1,\cdots,k \tag{4b. 2. 10}$$

で与えられる．任意に選んだある特定の $g(\theta)$ にたいして

$$P[\min_{\theta\in C} g(\theta), \max_{\theta\in C} g(\theta)]\geqslant 1-\alpha \tag{4b. 2. 11}$$

であることを注意する．したがって任意に選んだ特定の関数の信頼係数は $1-\alpha$ 以上である．(4b. 2. 9) において，F_α は F 統計量の上側 α 確率値である．

線形な母数の関数の同時信頼区間に応用できる他の方法はコーシー＝シュヴァルツの不等式 ((1e. 1. 4) を参照)

$$\mathbf{A}'\mathbf{B}^{-1}\mathbf{A}=\max_{\mathbf{U}}\frac{(\mathbf{U}'\mathbf{A})^2}{\mathbf{U}'\mathbf{B}\mathbf{U}} \tag{4b. 2. 12}$$

に関係している．ここで $\mathbf{U}$ と $\mathbf{A}$ は列ベクトルであり，$\mathbf{B}$ は正定符号行列である．

$\mathbf{A}=\mathbf{Z}-\theta$, $\mathbf{B}=\mathbf{D}$ として結果 (4b. 2. 12) を適用すると

$$\frac{(\mathbf{Z}-\theta)'\mathbf{D}^{-1}(\mathbf{Z}-\theta)}{ks^2}=\frac{1}{ks^2}\max_{\mathbf{U}}\frac{[\mathbf{U}'(\mathbf{Z}-\theta)]^2}{\mathbf{U}'\mathbf{D}\mathbf{U}} \tag{4b. 2. 13}$$

が得られる．したがって (4b. 2. 9) と (4b. 2. 13) とから，

$$P\left\{\max_{\mathbf{U}}\frac{|\mathbf{U}'(\mathbf{Z}-\theta)|}{\sqrt{\mathbf{U}'\mathbf{D}\mathbf{U}}}\leqslant s\sqrt{kF_\alpha}\right\}=1-\alpha,$$

すなわち

$$P(|\mathbf{U}'(\mathbf{Z}-\theta)|\leqslant s\sqrt{kF_\alpha\mathbf{U}'\mathbf{D}\mathbf{U}}\quad \text{すべての } \mathbf{U} \text{ にたいして})=1-\alpha$$

または　(4b. 2. 14)

$$P(\mathbf{U}'\theta\in\mathbf{U}'\mathbf{Z}\pm s\sqrt{kF_\alpha\mathbf{U}'\mathbf{D}\mathbf{U}}\quad \text{すべての } \mathbf{U} \text{ にたいして})=1-\alpha$$

が得られる．等式 (4b. 2. 14) は，形 $\mathbf{U}'\theta=\mathbf{U}'\mathbf{H}'\beta$ のすべての関数，すなわち与えられた母数の関数 $\mathbf{H}'\beta$ のすべての線形結合にたいする同時信頼区間を与える．また (4b. 2. 14) から，Scheffe (1959) によって指摘されているように，任意かつ特定の関数 $\mathbf{U}'\theta$ にたいして

$$P(\mathbf{U}'\theta\in\mathbf{U}'\mathbf{Z}\pm s\sqrt{kF_\alpha\mathbf{U}'\mathbf{D}\mathbf{U}})\geqslant 1-\alpha$$

が得られる．したがって信頼係数は $1-\alpha$ 以上である．

母数の2つの線形な関数の比にたいする信頼区間　$\theta_1=\mathbf{P}_1'\beta$, $\theta_2=\mathbf{P}_2'\beta$ を母数の2つの線形な関数とする．比 $\lambda=\theta_1/\theta_2$ にたいする信頼区間を求めよう．$\mathbf{P}_1'\beta, \mathbf{P}_2'\beta$ の l. s. 推定量を z_1, z_2 とし，これら推定量の分散，共分散を $C_{11}\sigma^2$, $C_{22}\sigma^2$, $C_{12}\sigma^2$ とする．関数 $\theta_1-\lambda\theta_2$ の l. s. 推定量は $z_1-\lambda z_2$ であって

$$E(z_1-\lambda z_2)=0,\qquad \lambda \text{が真の比であるとき}$$
$$V(z_1-\lambda z_2)=(C_{11}-2\lambda C_{12}+\lambda^2C_{22})\sigma^2$$

である．それゆえに

$$F=\frac{(z_1-\lambda z_2)^2}{s^2(C_{11}-2\lambda C_{12}+\lambda^2C_{22})}\sim F(1,n-r) \tag{4b. 2. 15}$$

であって

$$P[(z_1-\lambda z_2)^2-F_\alpha s^2(C_{11}-2\lambda C_{12}+\lambda^2C_{22})\leqslant 0]=1-\alpha \tag{4b. 2. 16}$$

である．(4b. 2. 16) の括弧内の不等式は λ の $1-\alpha$ 信頼域を与える．(4b. 2. 16) における式は λ についての2次式であるから，信頼域は，係数の性質に依存して，2次方程式の2根の間の区間であるかまたはこの区間の外部である．

4b. 3　制約のある模型

もとの模型 $(\mathbf{Y}, \mathbf{X}\beta, \sigma^2\mathbf{I})$ において，母数にある自然な制約式 $\mathbf{R}\beta=\eta$ があり，われわれは，その制約式によって値が特定されないような母数の関数についての推測をしたいものとする．この事態は理論的には困難はない．というのは，**4a. 9** に示したように，制約のあるもとの母数を制約のない新しい（より少数個の）母数におきかえることができるからである．ϕ で新しい母数をあらわすと，問題は $(\mathbf{Y}, \mathbf{B}\phi, \sigma^2\mathbf{I})$ の形のものとなる．もとの模型で推定可能である任意の母数の関数 $\mathbf{P}'\beta$ は，新しい模型で推定可能である母数の関数 $\mathbf{U}'\phi$ と同値であるというような関係があるから，制約のある場合の問題について特別な理論的議論は不必要である．しかしながら実際には，以下に示すように，すべての計算は母数の新しい集合を導入しないでなされる．残差平方和は

$$R_0^2=\min_{\phi}(\mathbf{Y}-\mathbf{B}\phi)'(\mathbf{Y}-\mathbf{B}\phi)=\min_{\mathbf{R}\beta=\eta}(\mathbf{Y}-\mathbf{X}\beta)'(\mathbf{Y}-\mathbf{X}\beta)$$

であって，その自由度は $n-f$ に等しい．ただし

$$f = \operatorname{rank}\mathbf{B} = \operatorname{rank}(\mathbf{X}' \vdots \mathbf{R}') - \operatorname{rank}\mathbf{R}'.$$

$\mathbf{H}'\boldsymbol{\beta} = \boldsymbol{\theta}_0$ を検定したいとしている階数 k の多重仮説であるとする．ただし $\mathbf{H}'\boldsymbol{\beta}$ における個々の関数のどの線形結合の値も，制約式 $\mathbf{R}\boldsymbol{\beta} = \boldsymbol{\eta}$ から導かれないものとする．$\boldsymbol{\phi}$ であらわした同値な仮説を $\mathbf{G}'\boldsymbol{\phi} = \boldsymbol{\xi}$ とする．そうすると

$$R_1^2 = \min_{\mathbf{G}'\boldsymbol{\phi}=\boldsymbol{\xi}} (\mathbf{Y} - \mathbf{B}\boldsymbol{\phi})'(\mathbf{Y} - \mathbf{B}\boldsymbol{\phi})$$
$$= \min_{\mathbf{H}'\boldsymbol{\beta}=\boldsymbol{\theta}_0, \mathbf{R}\boldsymbol{\beta}=\boldsymbol{\eta}} (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})$$

であって，検定基準は

$$\frac{R_1^2 - R_0^2}{k} \div \frac{R_0^2}{n-f}$$

となる．

もし $R_1^2 - R_0^2$ が，最初に $\mathbf{H}'\boldsymbol{\beta}$ の l. s. 推定量とその推定量の分散共分散行列を決定することにより計算されることが調べられるならば，R_0^2 を新しい母数に変換しないで求める方法はすでに 4a. 9 で説明している．

したがって，すべての計算は新しい母数を導入しないで行うことができるようである．

4c　単一標本の問題

4c. 1　検　定　基　準

$y_1, \cdots, y_n$ を平均 μ，分散 σ^2 の母集団からとられた観測値とする．次の計算をする．

$$R_0^2 = \min_{\mu} \sum (y_i - \mu)^2 = \sum y_i^2 - n\bar{y}^2, \qquad s^2 = \frac{R_0^2}{n-1}.$$

μ にたいする正規方程式は $\sum y = n\mu$ であるから，μ の l. s. 推定値は

$$\hat{\mu} = \frac{\sum y}{n} = \bar{y} \text{ であって，} \quad V(\hat{\mu}) = \frac{\sigma^2}{n}.$$

μ の推測をするための，(4b. 1. 1) で定義された t 統計量は

$$t = \frac{\sqrt{n}(\bar{y} - \mu)}{s} \sim S(n-1).$$

4c. 2　左右の大腿骨の非対称性（一対比較）

ある族の 36 個の骨格において，左右の大腿骨の長さの差（右大腿骨－左大腿骨）の平均は 2. 0234 であった．これら 36 個の差の修正された平方和 R_0^2 は 418. 6875 である．自由度 35 の分散の推定値 s^2 は 418. 6875/35 = 11. 9625 である．

両側検定　左右の大腿骨は平均として同じ長さであるか？　$\mu = 0$ とおいて，統計量

$$t = \frac{\sqrt{n}(\bar{x} - \mu)}{s} = \frac{\sqrt{36}(2.0234)}{\sqrt{11.9625}} = 3.5193$$

を用いる．左の測定値は右よりも大きい，あるいはその逆であるといったような事前の情報がないならば，両方向における t の大きい値は帰無仮説に反対する証拠となるであろう．t の観測値を，その符号を無視して，たいていの統計数値表（この章の終りの参考文献を参照）に与えられている $|t|$ の分布の選ばれた有意点と比べる．自由度 35 にたいしては，$|t|$ の 5％値は，R. M. M. 表からでは 2. 03 であり，この値は F. Y. 表に与えられている t の 5％値と同じである．観測された t は 5％水準で有意である．

片側検定　帰無仮説にたいする対立仮説は右大腿骨が左よりも長いということであることが事前にわかっているか，または検定の目的が，右大腿骨が左よりも長いという非対称性だけを判別することであるならば，t の符号は重要である．等しいという仮説は，t が t の上側 5％値――それは自由度 35 にたいしては R. M. M. 表から 1. 69 である（これは，F. Y. 表における t の 10％値と同じである）――をこえるときだけ，提起されている対立仮説に賛成して棄却される．もし対立仮説が左大腿骨はより大きい長さをもつということであるならば，$-t$ が上側 5％の有意値をこえなければいけない．前述のものでは，t は t の上側 5％値を確かにこえており，データは，右大腿骨は左よりも長いという提起された対立仮説と一致しているということを示している．任意の問題において，両側検定を用いるか片側検定を用いるかの決定はある論理的根拠に基づいてなされるべきである．

これらの検定は，2 つの族の平均値を比較したいが，観測値には，1 つの族の要素と他の族の要素との間に 1：1 の対応があるという場合に有用である．前の例では 2 つの測定値は同じ骨格からとられている．36 個の測定値の組は，平均の期待値が零であるような単一の族として取り扱われる 36 個の差のもとになっている．

一方，2 つの族が対応をもっていないかも知れない．たとえば，2 つの測定は同一の骨格でなされていない場合であり，この場合には群にたいして適用される分散分析の方法（4d を参照）が用いられるべきである．第 2 の検定では骨格による変動もまた考慮されるべきであり，したがって比較の精度は減少し，前述のもののような小さい差は，大標本においてさえも，検出できないかも知れない．大腿骨の長さの分散は約 400 であり，この場合には 36 個からなる 2 つの独立な族の平均の差の分散は

$$400\left(\frac{1}{36} + \frac{1}{36}\right) = 22.22$$

である．しかるに相関のある族における差の分散は 11. 9625 ÷ 36 = 0. 33 であり，これは非対称性についての相当に正確な評価を与える．この見方はどんな調査のときでも心に留めておくべきである．2 つの測定値を比較するときには，相関のある測定値の系列が得ら

れるように実験を計画するとよい．相関が高ければ高いほど利益は大きい．その関係は正であるべきである．そうでないならば相関のある組に基づく検定の有効さは減少する．

4d　1元分類のデータ

4d.1　検　定　基　準

未知の平均 $\mu_1, \cdots, \mu_k$ をもち，共通の未知の分散 σ^2 をもつ k 個の母集団から，大きさ $n_1, \cdots, n_k$ の k 個の標本があるとする．検定したい仮説は

$$\mu_1 = \mu_2 = \cdots = \mu_k \tag{4d.1.1}$$

である．$n_1 + \cdots + n_k = n$ 個の観測方程式を下に与える．仮説 (4d.1.1) の条件，つま

	第1番目の標本		…	第k番目の標本	
	変　量	期待値		変　量	期待値
	y_{11}	μ_1	…	y_{k1}	μ_k
	⋮	⋮	…	⋮	⋮
	y_{1n_1}	μ_1	…	y_{kn_k}	μ_k
合　計	T_1		…	T_k	

り共通の μ をもつという条件のもとでの $\sum\sum(y_{ij} - \mu_i)^2$ の最小値は

$$\min_{\mu} \sum\sum(y_{ij} - \mu)^2 = \sum\sum y_{ij}^2 - \frac{\left(\sum\sum y_{ij}\right)^2}{n} = R_1^2 \tag{4d.1.2}$$

であり，それは，すべての観測値の修正された総平方和である．制約のない場合の $\sum\sum(y_{ij} - \mu_i)^2$ の最小値は

$$\sum_i \min_{\mu_i} \sum_j (y_{ij} - \mu_i)^2 = \sum_{i=1}^{k}\left(\sum_j y_{ij}^2 - \frac{T_i^2}{n_i}\right) = R_0^2 \tag{4d.1.3}$$

である．したがって自由度 $k-1$ の仮説からの偏差による平方和は

$$R_1^2 - R_0^2 = \frac{T_1^2}{n_1} + \cdots + \frac{T_k^2}{n_k} - \frac{T^2}{n} \tag{4d.1.4}$$

$$T = T_1 + \cdots + T_k$$

表 4d.1α　分散分析，1元分類

	D.F.	S.S.
仮説からの偏差または標本間	$k-1$	$\sum\frac{T_i^2}{n_i} - \frac{T^2}{n}$
残差または標本内	*	*
合　　計	$n-1$	$\sum\sum y_{ij}^2 - \frac{T^2}{n}$

であり，標本の合計だけに関係している．実際には (4d.1.2) と (4d.1.4) を計算し，引き算によって R_0^2 の式 (4d.1.3) を求めるほうが容易である．計算の方法を表 4d.1α に示す．＊印のついているところの数値は引き算で求める．F 統計量は表 4d.1α から得られる平均平方を用いて作られる．

4d.2　1つの例題

次のデータは3つの族に属する142個の頭蓋骨の頭幅に関するものである．真の平均頭幅は3つの族で同じと考えることができるであろうか？

族	標本の大きさ	頭幅 合計	頭幅 平均
1	83	11,277	135.87
2	51	7,049	138.22
3	8	1,102	137.75
合　計	142	19,428	136.817

族間の平方和は

$$11{,}277 \times 135.87 + 7{,}049 \times 138.22 + 1{,}102 \times 137.75 - 19{,}428 \times 136.817 = 238.59$$

である．総平方和は 4,616.64 であった．分散分析を表 4d.2 に示す．

表 4d.2　分散分析

	D.F.	S.S.	M.S.	F
族　間	2	238.59	119.29	3.79
族　内	139	4,378.05	31.50	
合　計	141	4,616.64		

自由度2，139の分散比3.79は5％水準で有意であり，平均値に差があることを示している（F の上側有意値にたいしては R.M.M. 表を参照）．さて，平均値を大きい順に書くことによりこれらの差の性質を調べることができる．

族	2	3	1
平　均	138.22	137.75	135.87

(2) と (3) の間の観測された不等性は不確かであるように思える．なぜならばそれらの値が近いからである．このことを調べるために，5％の限界差

$$\left[\left(\frac{1}{n_2} + \frac{1}{n_3}\right)\hat{\sigma}^2\right]^{\frac{1}{2}} t_\alpha$$

を計算してもよい．ここで n_2, n_3 は標本の大きさであり，$\hat{\sigma}^2$ は σ^2 の推定値であって，そ

の値は表4d. 2から31. 50で自由度は139である．またt_αは自由度139の$|t|$のα水準の有意値である．数値を入れると，自由度139の$|t|$の5％値は1. 96であるから，棄却差は

$$\left[\left(\frac{1}{51}+\frac{1}{8}\right)31.50\right]^{\frac{1}{2}}t_\alpha = 2.13t_\alpha = 2.13\times 1.96 = 4.17$$

である．この棄却値は観測された差 138. 22 − 137. 75 = 0. 47 よりも相当に大きいので，平均値における全体の差は族(1)が(2)，(3)と異なっているせいであるとみてよい．一般に，すべての可能な組にたいして棄却差を上に述べたようにして計算し，観測された差と比べてもよい．このような詳細な比較に基づいて，平均値に関していま比較しようとしている群の順序を決めたり，そのなかでは順序がやや不確かであるというような部分集合をきめたりすることができる．いまの例では，族(1)は最後の位置をとるが，(2)と(3)の間の順序は不確かであると結論してよい．

実際の場面では，順序だけでなく，差の大きさもまた重要である．データの分散分析，平均値の表示，個々の差の調査は，差の大きさの上手な解釈を可能にする．

4e　2元分類のデータ

4e. 1　各級で1つの観測値の場合

pq個の観測値があり，おのおのは2つの種類のカテゴリーで規定されるものとする．観測値は表4e. 1αのように並べられる．

表 4e. 1α

種類 A	種類 B: B_1	B_2	$\cdots$	B_p	合計
A_1	y_{11}	y_{12}	$\cdots$	y_{1p}	$y_{1\cdot}$
A_2	y_{21}	y_{22}	$\cdots$	y_{2p}	$y_{2\cdot}$
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
A_q	y_{q1}	y_{q2}	$\cdots$	y_{qp}	$y_{q\cdot}$
合計	$y_{\cdot 1}$	$y_{\cdot 2}$		$y_{\cdot p}$	$y_{\cdot\cdot}$

Aの異なったカテゴリーは小麦の品種であり，Bは天火であり，そして測定値は焼いたパンの品質かも知れない．Aは機械の異なった型をあらわし，Bは何人かの作業者をあらわしそして測定値は生産の効率かも知れない．または，Aは作物に与えるある有機肥料の異なった水準でBはある無機肥料の異なった水準かも知れない．これらすべての場合にお

いて，目的は，A, Bのカテゴリーの最適な組合せを決定することであるか，または一方の因子の適当に定義されたカテゴリーにたいして1つの因子のカテゴリーの間の差を研究することである．

$E(y_{ij}) = \alpha_i + \beta_j$，$V(y_{ij}) = \sigma^2$とする．観測値は$n = pq$個，母数は$p+q$個あるが，観測方程式の行列のランクはただの$p+q-1$である(11頁の問題3を参照)．$\sum\sum(y_{ij}-\alpha_i-\beta_j)^2$を最小にすることによって得られる正規方程式は

$$p\alpha_i + \sum\beta_j = y_{i\cdot} \qquad i = 1, \cdots, q$$

$$\sum\alpha_i + q\beta_j = y_{\cdot j} \qquad j = 1, \cdots, p$$

であり，これの解は一意には定まらないが，1つの解は

$$\hat{\alpha}_i = \frac{y_{i\cdot}}{p}, \qquad \hat{\beta}_j = \frac{y_{\cdot j}}{q} - \frac{y_{\cdot\cdot}}{pq}$$

であることがわかる．しかしこれからあとの計算には1つの解だけが必要である．

残差平方和にたいする公式(4a. 5. 1)を用いると

$$R_0^2 = \sum\sum y_{ij}^2 - \sum\hat{\alpha}_i y_{i\cdot} - \sum\hat{\beta}_j y_{\cdot j}$$

$$= \left(\sum\sum y_{ij}^2 - \frac{y_{\cdot\cdot}^2}{pq}\right) - \left(\sum\frac{y_{i\cdot}^2}{p} - \frac{y_{\cdot\cdot}^2}{pq}\right) - \left(\sum\frac{y_{\cdot j}^2}{q} - \frac{y_{\cdot\cdot}^2}{pq}\right) \qquad (4e.\ 1.\ 1)$$

であり，その自由度は$n - r = pq - (p+q-1) = (p-1)(q-1)$であることがわかる．仮説

$$\alpha_1 = \cdots = \alpha_q \qquad (4e.\ 1.\ 2)$$

を検定したいとしよう．この場合には，共通のαを用いて$E(y_{ij}) = \alpha + \beta_j$である．正規方程式を書き，$\alpha, \beta_1, \cdots, \beta_p$にたいして解き，残差平方和にたいする公式を適用すると

$$R_1^2 = \min\sum\sum(y_{ij}-\alpha-\beta_j)^2$$

$$= \left(\sum\sum y_{ij}^2 - \frac{y_{\cdot\cdot}^2}{pq}\right) - \left(\sum\frac{y_{\cdot j}^2}{q} - \frac{y_{\cdot\cdot}^2}{pq}\right)$$

であることがわかる．したがって(4e. 1. 2)を検定するための平方和は

$$R_1^2 - R_0^2 = \sum\frac{y_{i\cdot}^2}{p} - \frac{y_{\cdot\cdot}^2}{pq} \qquad (4e.\ 1.\ 3)$$

であり，自由度は$q-1$である．同様にして仮説$\beta_1 = \cdots = \beta_p$を検定するための平方和も決定することができる．計算の方法を表4e. 1βに与える．仮説(4e. 1. 2)を検定するための分散比はV_A/V_eである．

表4e. 1βにおける分散分析の成分は，データについてのもっと一般的な仮定のもとで，1つの因子のカテゴリー間の差の検定をするのにも有用である．具体的な例として，因子Aは機械の異なった種類，因子Bは作業者の異なった種類をあらわす場合をとりあげよう．もし仮定$E(y_{ij}) = \alpha_i + \beta_j$が正しいならば，機械の差は各作業者にたいして同じで

表 4e. 1β　分散分析，2元分類

	D. F.	S. S.	M. S.
A種間	$q-1$	$S_A=\frac{1}{p}\sum y_{i\cdot}^2-\frac{1}{pq}y_{\cdot\cdot}^2$	V_A
B種間	$p-1$	$S_B=\frac{1}{q}\sum y_{\cdot j}^2-\frac{1}{pq}y_{\cdot\cdot}^2$	V_B
残差（交互作用）$A\times B$	$(p-1)(q-1)$	$S_{AB}=$ * （引き算による）	V_e
	$pq-1$	$S=\sum\sum y_{ij}^2-\frac{1}{pq}y_{\cdot\cdot}^2$	

あり，それは母数 α_i における差で測られる．一方，機械間の差は各作業者にたいして同じではないかも知れないが，実験に含まれている p 人の作業者が標本であるような作業者の広い集団（母集団）に関して，平均として機械間に差があるかどうかの問題を調べたい．しかしながら，1人の作業者の q 個の異なった機械にたいする q 個の観測値は対称な分布をもつ，つまり，等しい平均，等しい分散をもち，観測値のすべての対にたいして同じ相関をもつ，という広い帰無仮説を考えよう．または他の言葉で言うと，名前のちがいは別として，機械は根本的に同じとするのである．これらの条件のもとで，q 個の機械の平均値が作業者の母集団にたいして同じであるならば，V_A/V_e は同じ分散比分布をもつ［自由度は $q-1$ と $(p-1)(q-1)$］ことが (3c. 2. 10) で示されている．しかしながら検定 V_A/V_e は，機械の平均値の差を検出することにだけ有効であり，考えている広い帰無仮説の他のズレにたいしては有効ではない．かくして検定 V_A/V_e は仮定したモデルに依存して異なった解釈を許す．

非加法性にたいするテューキー（1949）の検定　加法性からのズレは，積の項を導入して

$$E(y_{ij})=\mu+\alpha_i+\beta_j+\lambda\alpha_i\beta_j \qquad \sum\alpha_i=0,\quad \sum\beta_j=0 \tag{4e. 1. 4}$$

のように，y_{ij} の期待値を書くことによって規定されるものとする．もし仮説 $\lambda=0$ の検定をつくることができるならば，それは非加法性の検定，特にタイプ (4e. 1. 4) のズレを検出するための検定として役立つであろう．$\lambda\neq 0$ のとき期待値は線形ではないから，最小2乗法の理論は適用できない．

しかし $\lambda=0$ であろうとあるまいと

$$E(\bar{y}_{i\cdot}-\bar{y}_{\cdot\cdot})=\alpha_i,\qquad E(\bar{y}_{\cdot j}-\bar{y}_{\cdot\cdot})=\beta_j,\qquad E(\bar{y}_{\cdot\cdot})=\mu$$

であることがわかるから，α_i,β_j,μ の不偏推定量は

$$\alpha_i^*=\bar{y}_{i\cdot}-\bar{y}_{\cdot\cdot},\quad \beta_j^*=\bar{y}_{\cdot j}-\bar{y}_{\cdot\cdot},\quad \mu^*=\bar{y}_{\cdot\cdot}$$

で与えられる．さらに

$$E(y_{ij}-\bar{y}_{i\cdot}-\bar{y}_{\cdot j}+\bar{y}_{\cdot\cdot})=\lambda\alpha_i\beta_j$$

であって，α_i と β_j がわかっているときには，(4e. 1. 4) に最小2乗法の理論を用いて，λ の推定値を

$$\hat{\lambda}=\frac{\sum\sum\alpha_i\beta_j(y_{ij}-\alpha_i-\beta_j-\bar{y}_{\cdot\cdot})}{\sum\sum\alpha_i^2\beta_j^2}$$

と書いてもよい．推定値 α_i^*,β_j^* を代入すると，わかった量を用いた λ の推定値

$$\lambda^*=\frac{\sum\sum(\bar{y}_{i\cdot}-\bar{y}_{\cdot\cdot})(\bar{y}_{\cdot j}-\bar{y}_{\cdot\cdot})(y_{ij}-\bar{y}_{i\cdot}-\bar{y}_{\cdot j}+\bar{y}_{\cdot\cdot})}{\sum(\bar{y}_{i\cdot}-\bar{y}_{\cdot\cdot})^2\sum(\bar{y}_{\cdot j}-\bar{y}_{\cdot\cdot})^2}$$

を得る．これは簡単にすると

$$\frac{pq\sum\sum(\bar{y}_{i\cdot}-\bar{y}_{\cdot\cdot})(\bar{y}_{\cdot j}-\bar{y}_{\cdot\cdot})y_{ij}}{S_AS_B} \tag{4e. 1. 5}$$

となり，α_i^*,β_j^* が与えられたときのこれの分散は

$$\frac{p^2q^2\sigma^2\sum\sum(\bar{y}_{i\cdot}-\bar{y}_{\cdot\cdot})^2(\bar{y}_{\cdot j}-\bar{y}_{\cdot\cdot})^2}{S_A^2S_B^2}=\frac{pq\sigma^2}{S_AS_B} \tag{4e. 1. 6}$$

である．したがって，$\lambda=0$ であるならば，$E(\lambda^*|\alpha_i^*,\beta_j^*$，すべての i,j にたいして$)=0$ であって，

$$\frac{(4e.\ 1.\ 5)^2}{(4e.\ 1.\ 6)}=\frac{pq\,[\sum\sum(\bar{y}_{i\cdot}-\bar{y}_{\cdot\cdot})(\bar{y}_{\cdot j}-\bar{y}_{\cdot\cdot})y_{ij}]^2}{S_AS_B\sigma^2}\sim\chi^2(1). \tag{4e. 1. 7}$$

しかし

$$\frac{\sum\sum(y_{ij}-\bar{y}_{i\cdot}-\bar{y}_{\cdot j}+\bar{y}_{\cdot\cdot})^2}{\sigma^2}=\frac{S_{AB}^2}{\sigma^2}\sim\chi^2(\overline{p-1}\,\overline{q-1}) \tag{4e. 1. 8}$$

であって，［(4e. 1. 7) における y_{ij} は $(y_{ij}-\bar{y}_{i\cdot}-\bar{y}_{\cdot j}+\bar{y}_{\cdot\cdot})$ と書かれるから］，(4e. 1. 8)–(4e. 1. 7) は非負であることがわかる．したがって［(iv)，3b. 4］を適用すると

$$\frac{\chi^2(1)}{\chi^2(\overline{p-1}\,\overline{q-1})-\chi^2(1)}\cdot\frac{(p-1)(q-1)-1}{1} \tag{4e. 1. 8}$$

は自由度が1と $(p-1)(q-1)-1$ の正確な F 分布をする．〔式番 (4e. 1. 8) がダブっているが，原著に従った.〕 さて，表 4e. 1γ のように，非加法性にたいする成分を含んでいる完全な分散分析を書くことができる．ここで，S_A,S_B,S は表 4e. 1β の中のものと同じであり，S_N は (4e. 1. 7) に定義されているように

$$S_N=pq[\sum\sum(\bar{y}_{i\cdot}-\bar{y}_{\cdot\cdot})(\bar{y}_{\cdot j}-\bar{y}_{\cdot\cdot})y_{ij}]^2/S_AS_B$$

である．非加法性にたいする分散比は V_N/V_R であって，自由度1と $(p-1)(q-1)-1$ をもつ．

表 4e. 1τ　2元データの分散分析

	D. F.	S. S.	M. S.
A	$q-1$	S_A	V_A
B	$p-1$	S_B	V_B
非加法性	1	S_N	V_N
残　差	$(p-1)(q-1)-1$	* (引き算による)	V_R
合　計	$pq-1$	S	

テューキーの検定 (4e. 1. 8) のここで与えた誘導法と解釈は，ガウス＝マルコフ模型 $(\mathbf{Y}, \mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I})$ においてモデルの検定に関する次の一般化を提起する．

実際のモデルは

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{F}\boldsymbol{\lambda} + \boldsymbol{\varepsilon} \tag{4e. 1. 9}$$

であるとする．ここで $\mathbf{F}$ は，(i, j) 要素が $\mathbf{X}\boldsymbol{\beta}$ の与えられた関数 $F_{ij}(\mathbf{X}\boldsymbol{\beta})$ であるような $n \times k$ の行列である．われわれは仮説 $\boldsymbol{\lambda} = \mathbf{0}$ を検定したい．

模型 $(\mathbf{Y}, \mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I})$ からの $\boldsymbol{\beta}$ の最小2乗推定量を $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^-\mathbf{X}'\mathbf{Y}$ とする．さらに

$$\mathbf{d} = \mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}$$

$$\mathbf{M} = (\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^-\mathbf{X}')\hat{\mathbf{F}} \tag{4e. 1. 10}$$

とする．ただし $\hat{\mathbf{F}}$ の (i, j) 要素は $F_{ij}(\mathbf{X}\hat{\boldsymbol{\beta}})$ である．模型 $(\mathbf{d}, \mathbf{M}\boldsymbol{\lambda}, \sigma^2\mathbf{I})$ を（形式的に）考え，**4b** の最小2乗法の理論を（形式的に）用い，仮説 $\boldsymbol{\lambda} = \mathbf{0}$ を検定するための統計量 (4b. 2. 8) を求めよう．

4b. 2 の記号を用いると

$$R_1^2 = \mathbf{d}'\mathbf{d}, \qquad R_0^2 = \mathbf{d}'(\mathbf{I} - \mathbf{M}(\mathbf{M}'\mathbf{M})^-\mathbf{M}')\mathbf{d}$$

であるから，求める統計量は

$$\frac{a}{b}\cdot\frac{R_1^2 - R_0^2}{R_0^2} = \frac{a}{b}\cdot\frac{\mathbf{d}'\mathbf{M}(\mathbf{M}'\mathbf{M})^-\mathbf{M}'\mathbf{d}}{\mathbf{d}'(\mathbf{I} - \mathbf{M}(\mathbf{M}'\mathbf{M})^-\mathbf{M}')\mathbf{d}}$$

$$= \frac{a}{b}\cdot\frac{\mathbf{Y}'\mathbf{M}(\mathbf{M}'\mathbf{M})^-\mathbf{M}'\mathbf{Y}}{\mathbf{Y}'[\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^-\mathbf{X}' - \mathbf{M}(\mathbf{M}'\mathbf{M})^-\mathbf{M}']\mathbf{Y}} \tag{4e. 1. 11}$$

である．統計量 (4e. 1. 11) は $F(b, a)$ に従って分布する．ただし

$$b = R(\mathbf{M}), \qquad a = n - R(\mathbf{X}) - b \tag{4e. 1. 12}$$

で，b は確率1で定数であるとする．$b \neq k$ ならば，$\boldsymbol{\lambda}$ のある線形関数が零であるという部分仮説を検定していることになる．

(4e. 1. 11) の証明は公式 (4e. 1. 7) と (4e. 1. 8) において用いたものと同じである．〔ここでの (4e. 1. 8) は，2つある (4e. 1. 8) 式の後者を指す．〕しかしながら，結果 (4e. 1. 11) は，$\hat{\mathbf{F}}$ と $\mathbf{d}$ が独立であることに注目すると，一般定理〔(iv)，**3b. 5**〕を適用してもっ

と簡単に導かれる．Millikin and Graybill (1970) は検定 (4e. 1. 11) を得ているが，彼等の場合には自由度 a, b は異なって定義されている．

4e. 2　各級で2つ以上の等しい数の観測値がある場合

4e. 1 において，仮説 $E(y_{ij}) = \alpha_i + \beta_j$，つまり種類 A（たとえば機械）と B（たとえば作業者）による効果が加法的であるという仮説の不完全な検定を考えた．しかしながらこれは，各級で2つ以上の観測値がある場合には完全な検定ができる．級 (i, j) で c 個の観測値があるならば，それらは

$$y_{ij1}, \quad \cdots, \quad y_{ijc}$$

$$E(y_{ijk}) = \mu_{ij}, \qquad V(y_{ijk}) = \sigma^2 \tag{4e. 2. 1}$$

であらわされる．検定したい仮説は

$$\mu_{ij} = \alpha_i + \beta_j \tag{4e. 2. 2}$$

である．加法的であるという仮説 (4e. 2. 2) が正しくないとき，因子 A と B との間には交互作用があると呼ばれる．交互作用が存在するときには，A 種類間の差を B 種類の特定のものに関係なく規定することはできない．

$$y_{ij} = \sum_k y_{ijk}, \qquad y_{..} = \sum_i \sum_j y_{ij},$$

$$y_{i.} = \sum_j y_{ij}, \qquad y_{.j} = \sum_i y_{ij}$$

とする．そうすると模型 (4e. 2. 1) のもとでは

$$R_0^2 = \left(\sum\sum\sum y_{ijk}^2 - \frac{y_{..}^2}{pqc}\right) - \left(\sum\sum \frac{y_{ij}^2}{c} - \frac{y_{..}^2}{pqc}\right)$$

となり，仮説 (4e. 2. 2) のもとでは，(4e. 1. 1) と同じように

$$R_1^2 = \left(\sum\sum\sum y_{ijk}^2 - \frac{y_{..}^2}{pqc}\right) - \left(\sum \frac{y_{i.}^2}{pc} - \frac{y_{..}^2}{pqc}\right) - \left(\sum \frac{y_{.j}^2}{qc} - \frac{y_{..}^2}{pqc}\right)$$

となる．交互作用の平方和は $R_1^2 - R_0^2$ で，その自由度は $(p-1)(q-1)$ である．計算の方法を表 4e. 2α に与える．

表 4e. 2α のような2元データの分散分析から，いろいろな仮説の検定をすることができる．

1. 交互作用は有意であるか？　すでに示しているように，この仮説を検定するための分散比は V_{AB}/V_e である．
2. それぞれの作業者にたいして機械間に差はあるかも知れないが，実験に従事した作業者（B 種類）を平均したとき，機械（A 種類）間に差があるか？　組合せ (A_iB_j) の効果を μ_{ij} とすると，検定したい仮説は

$$\mu_{1.} = \mu_{2.} = \cdots = \mu_{q.}$$

表 4e. 2α　分散分析，2元分類

	D. F.	S. S.	M. S.
機械間(A)	$q-1$	$\frac{1}{c}\frac{1}{p}\sum_1^p y_{i\cdot}^2 - \frac{1}{cpq}y_{\cdot\cdot}^2$	V_A
作業者間(B)	$p-1$	$\frac{1}{c}\frac{1}{q}\sum_1^q y_{\cdot j}^2 - \frac{1}{cpq}y_{\cdot\cdot}^2$	V_B
交互作用($A\times B$)	$(p-1)(q-1)$	* (引き算による)	V_{AB}
pq 個の級間	$pq-1$	$\frac{1}{c}\sum\sum y_{ij}^2 - \frac{1}{cpq}y_{\cdot\cdot}^2$	V
残　差	$pq(c-1)$	* (引き算による)	V_e
合　計	$pqc-1$	$\sum\sum\sum y_{ijk}^2 - \frac{1}{cpq}y_{\cdot\cdot}^2$	

である．これにたいする適切な分散比は V_A/V_e である．

3.　同様にして，仮説

$$\mu_{\cdot 1} = \mu_{\cdot 2} = \cdots = \mu_{\cdot p}$$

にたいする適切な分散比は V_B/V_e である．

4.　実験に従事した作業者が標本であるような作業者の広い集団（母集団）の上で平均したとき，機械間に差があるか？　この問題は **4e. 1** において議論したものと類似している．適切な分散比は V_A/V_{AB} である，すなわち，加法性が成立しようとしまいと，交互作用平均平方が正しい分母である．比 V_B/V_{AB} は B 種類についての同様な仮説を検定するのに使われる．

(2) と (4) で考えた仮説の間のちがいは理解しておくべきである．どちらの場合においても，交互作用は存在してもよい．(2) では，実験に従事した特定の作業者間の平均的な差に興味があるが，(4) では，差は，実験に含まれた作業者の全体ではなくて，それがただの小さい標本であるような作業者の母集団に関してである．交互作用がなければ (2) と (4) は同じである．

4e. 3　級における繰り返し数が等しくない場合

次の記号を用いる：

y_{ij} = (i, j) 番目の級における観測値の合計

$\bar{y}_{ij}$ = (i, j) 番目の級における平均

$y_{\cdot j} = \sum_i y_{ij}$ (j 番目の列の合計)

$y_{i\cdot} = \sum_j y_{ij}$ (i 番目の行の合計)

$y_{\cdot\cdot}$ = 観測値の合計

$\bar{y}_{\cdot\cdot}$ = 全観測値の平均

$n_{ij}, n_{\cdot j}, n_{i\cdot}, n_{\cdot\cdot}$ はそれぞれ，(i, j) 番目の級，j 番目の列，i 番目の行，全部の級における観測値の数である．

表 4e. 3α のデータは，3人の観測者によって3つの異なった層から発掘された頭蓋骨の鼻の高さの平均値と合計をあらわす．層や観測者の差の検定をしたい．

この種の問題では，種々の平方和を計算するのに，表 4e. 3α のように数字をならべると便利である．解析は3段階でなされる．自由度309の，すべての観測値の（修正された）総平方和は 5398. 4206 であることがわかっている．これから先の計算にたいしては，表 4e. 3α のなかの数値で十分である．

級間平方和の計算　自由度 $pq-1=8$ の級間平方和は，$\sum\sum y_{ij}\bar{y}_{ij} - y_{\cdot\cdot}\bar{y}_{\cdot\cdot} =$ 931. 5204 である．

表 4e. 3α　平均値と合計

観測者		層 S_1	S_2	S_3	$y_{i\cdot}$	$\bar{y}_{i\cdot}$
O_1	y_{11}	1071. 00	1572. 48	913. 50	3556. 98	
	$\bar{y}_{11}$	51. 00	49. 14	50. 75		50. 098
	n_{11}	(21)	(32)	(18)	(71)	
O_2		1966. 86	2315. 40	1721. 52	6003. 78	
		46. 83	45. 40	47. 82		46. 541
		(42)	(51)	(36)	(129)	
O_3		1219. 00	2091. 60	1849. 20	5159. 80	
		48. 76	46. 48	46. 23		46. 907
		(25)	(45)	(40)	(110)	
$y_{\cdot j}$		4256. 86	5979. 48	4484. 22	14720. 56	
$\bar{y}_{\cdot j}$		48. 373	46. 715	47. 704	$= y_{\cdot\cdot}$	47. 4857
		(88)	(128)	(94)	(310)	$= \bar{y}_{\cdot\cdot}$

主効果と交互作用の計算　(i, j) 番目の級の観測値の期待値は $\alpha_i + \beta_j$ であるとする．これは i 番目の行と j 番目の列の効果の加法性をあらわす．

$$\sum\sum n_{ij}(\bar{y}_{ij} - \alpha_i - \beta_j)^2$$

を最小にすることによって得られる正規方程式は次のようになる．ここで対称性の理由により対角線より下の数値は省略している．

観測者			層				縁の合計
α_1	α_2	α_3	β_1	β_2	β_3		
71	·	·	21	32	18	=	3556.98
	129	·	42	51	36	=	6003.78
		110	25	45	40	=	5159.80
			88	·	·	=	4256.86
				128	·	=	5979.48
					94	=	4484.22

これらの方程式を書く方法は簡単である．任意の縁（ふち）の合計，たとえば層の上に 21, 32, 18 として分布している観測者1の71個の観測値に基づく3556.98から出発する．これが最初の方程式を与える．AとBの種類に対応して6つの縁の合計があり，これが6つの方程式を与える．分散分析表をつくるのにこれらの方程式を実際に解く必要はない．$\beta_1, \beta_2, \beta_3$ にたいする終りの方程式の組から $\alpha_1, \alpha_2, \alpha_3$ を消去しよう．その規則はまず α_i または β_j にたいする方程式を書くことである．それは常にたくさんのほかの未知数を含んでおり，方程式の残りの組から対応する定数を消去する．β_1 にたいする方程式から $\alpha_1, \alpha_2, \alpha_3$ を消去するには，最初の方程式に 21/71，第2の方程式に 42/129，第3の方程式に 25/110 をそれぞれかけ，それらの合計を β_1 にたいする方程式から引かねばならない．同様に，第1，第2，第3 の方程式に 32/71，51/129，45/110 をそれぞれかけ，それらの合計を β_2 にたいする方程式から引くということをする．そうすると次の方程式が得られる．

β_1	β_2	β_3		
62.432498	−46.296717	……	=	77.394630
	64.844662	……	=	−108.080590

β_3 の係数と最後の方程式は計算する必要はない．β_i に関するこの方程式を最後の β はゼロに等しいとおいておく．いまの問題ではたった2つの方程式であり $\hat{\beta}_1 = 0.559250$, $\hat{\beta}_2 = -1.170335$ が得られる．Bの主効果にたいする平方和は，

$$\hat{\beta}_1(77.394630) + \hat{\beta}_2(-108.080590) = 171.8919$$

として簡単に求められる．ここで 77.394630 などの数値は β_i に関する縮約された方程式における右辺の要素である．分散分析表をつくるのに α_i について解く必要はない．しかし，さらに2つの式を計算しなければいけない．もしB種類を無視するならば，データは単にAに関する1元配置となり，自由度2のAによる平方和（Bを無視）はAの縁の合計から求められる．

$$\sum y_{i\cdot}\bar{y}_{i\cdot} - y_{\cdot\cdot}\bar{y}_{\cdot\cdot} = 3556.98 \times 50.098 + \cdots - 14720.56 \times 47.4857$$
$$= 634.1516$$

同様にして，Bによる平方和（Aを無視）を求める．あとの計算は分散分析表 4e. 3β に説明してある．

表 4e. 3β　2元データの分散分析（級における繰り返し数が等しくない場合）

根源	D. F.	S. S.		S. S.	D. F.	根源
層 B（Aを無視）	2	147.6319†		634.1516†	2	観測者 A（Bを無視）
観測者	2	658.4116*		171.8919†	2	層
交互作用	4	125.4769	←	125.4769*	4	交互作用
級間	8	931.5204†	→	931.5204	8	
級内	301	4466.9002*				
合計	309	5398.4206†				

† 本文で説明しているようにして直接計算する．
* 引き算で求める．→は，次の計算のため，1つの位置から他への数値（直接計算されたものまたは引き算で求められたもの）の移動をあらわす．

自由度4，301の交互作用の検定は

$$F = \frac{125.4769}{4} \div \frac{4466.9002}{301} = 2.11$$

であり，これは自由度4，301のFの5％値よりも小さいけれども，この値はやや大きい．この問題において交互作用の存在は，観測者の差は存在するが，その差は層にたいして一致していなく，したがって意味のある結論がデータからひき出せないという意味においておもしろくないものである．一方，観測者が層から標本を厳密な意味においてランダムにとらないということにより――このことは，頭蓋骨が異なった観測者により層の異なった場所からとられるならば，（困難な実際の条件のもとでは）ほんとうかも知れない――交互作用はいくらか現われるであろう．交互作用平均平方は級内平均平方よりも大きいから，観測者間，層間の差を検定するのに前者を用いる．

観測者にたいする自由度2，4の分散比は

$$\frac{658.4116}{2} \div 31.3692 = 10.49$$

であり，これは5％水準で有意である．層にたいする自由度2，4の分散比は

$$\frac{171.8919}{2} \div 31.3692 = 2.74$$

であり，これはさほど大きくはない（5％水準で有意ではない）．

観測者を無視し，層間の差を調べたとしよう．そのときの解析は

	D. F.	M. S.	比
層間	2	73.8159	4.31
層内	307	17.1035	

となり，比4.31は有意であって，層間の差は存在するという結論が得られるであろう．周到な解析は，観測者の差は重要であり，したがって人のちがった調査者の結果を一緒にする際にはある注意が必要であるということを知らせる．観測者による差を修正すると層間の差はさほど顕著ではない．

4f　2元データにたいする一般模型と分散成分

4f.1　一　般　模　型

2元データの分散分析から，観測値にたいするいろいろな模型のもとでいろいろな推論をすることができるということを，すでに4eにおいて調べた．ここでは，表4e.2αの分散分析が，有用な観測値の縮約を与える非常に一般的な模型を考えよう．pq 個の各級で c 個の観測値をもつ pqc 個の観測値を

$$y_{ijk},\quad i=1,\cdots,q\,;\,j=1,\cdots,p\,;\,k=1,\cdots,c \tag{4f.1.1}$$

であらわし，それらの期待値，分散，共分散は次のようであるとする：

$$E(y_{ijk})=\mu_{ij},\quad V(y_{ijk})=\sigma^2,\quad \mathrm{cov}(y_{ijk},y_{ijr})=\rho_0\sigma^2$$
$$\mathrm{cov}(y_{ijk},y_{imr})=\rho_1\sigma^2,\qquad \mathrm{cov}(y_{ijk},y_{mjr})=\rho_2\sigma^2$$
$$\mathrm{cov}(y_{ijk},y_{smr})=\rho_3\sigma^2 \tag{4f.1.2}$$

したがって任意の2つの観測値の間の相関（それは正，負の値をとり得る）は共通の添字の性質だけに関係する．

4f.2　分散成分模型

分散成分模型として知られている特別な場合を考えよう．それは，分散共分散構造（4f.1.2）をもつ型（4f.1.1）の観測値に帰着される．$f_{\alpha\beta}$ を1組の確率変数 (α,β) の関数であるとする．

$$\underset{\alpha}{E}\underset{\beta}{E}(f_{\alpha\beta})=f_{\cdot\cdot},\qquad \underset{\beta}{E}(f_{\alpha\beta}|\alpha)=f_{\alpha\cdot},\qquad \underset{\alpha}{E}(f_{\alpha\beta}|\beta)=f_{\cdot\beta} \tag{4f.2.1}$$

と定義する．そうすると次式が成立することが容易にわかる．

$$EE(f_{\alpha\beta}-f_{\cdot\cdot})^2=E(f_{\alpha\cdot}-f_{\cdot\cdot})^2+E(f_{\cdot\beta}-f_{\cdot\cdot})^2+EE(f_{\alpha\beta}-f_{\alpha\cdot}-f_{\cdot\beta}+f_{\cdot\cdot})^2$$
$$\sigma_f^2=\sigma_A^2\qquad+\qquad\sigma_B^2\qquad+\qquad\sigma_{AB}^2 \tag{4f.2.2}$$

変量αは因子 A のカテゴリーをあらわし，βは因子 B のカテゴリーをあらわすと考えてよく，$f_{\alpha\beta}$ は α,β の選定から得られる産出高である．かくして σ_A^2,σ_B^2 は，A,B のカテゴリーの選定にたいする与えられた標本抽出機構のもとでの，A のカテゴリーによる産出高の分散，B のカテゴリーによる産出高の分散と解釈してよい．（4f.2.2）における最後

の成分 σ_{AB}^2 は，因子 A と B との間の交互作用，すなわち $f_{\alpha\beta}$ が1つはαだけに関係し他の1つはβだけに関係するという2つの関数の和としてあらわされない状態，をあらわす．もし $f_{\alpha\beta}=f_{\cdot\cdot}+f_{\alpha\cdot}+f_{\cdot\beta}$ のような分解が真であるならば，A と B の効果は加法的であると呼ばれ，成分 σ_{AB}^2 はゼロとなる．

次の方法（標本抽出機構）で確率変数をつくろう．まず α,β の独立な標本 $\alpha_1,\cdots,\alpha_q$ と $\beta_1,\cdots,\beta_p$ を選び，組合せ (α_i,β_j) のおのおのに c 個の変量

$$y_{ijk}=f_{\alpha_i\beta_j}+e_{ijk},\qquad k=1,\cdots,c$$

ただし

$$E(e_{ijk}|\alpha_i,\beta_j)=0,\qquad V(e_{ijk}|\alpha_i,\beta_j)=\sigma_e^2$$
$$\mathrm{cov}(e_{ijk},e_{ijm}|\alpha_i,\beta_j)=0 \tag{4f.2.3}$$

をつくる．成分 e_{ijk} は $f_{\alpha_i\beta_j}$ の繰り返し測定における独立な誤差と考えてよい．さて（4f.2.2）で定義した成分 $\sigma_A^2,\sigma_B^2,\sigma_{AB}^2$ を用いると，α,β についての独立な選定によってつくられる (α,β) の特定の分布に関して，

$$E(y_{ijk})=f_{\cdot\cdot},\qquad V(y_{ijk})=\sigma_f^2+\sigma_e^2=\sigma^2$$
$$\mathrm{cov}(y_{ijk},y_{ijm})=\sigma_f^2=\rho_0\sigma^2,\qquad \mathrm{cov}(y_{ijk},y_{imr})=\sigma_A^2=\rho_1\sigma^2,$$
$$\mathrm{cov}(y_{ijk},y_{mjr})=\sigma_B^2=\rho_2\sigma^2,\qquad \mathrm{cov}(y_{ijk},y_{smr})=0=\rho_3\sigma^2,$$
$$\sigma_f^2=\sigma_A^2+\sigma_B^2+\sigma_{AB}^2 \tag{4f.2.4}$$

であることがわかる．かくして，（4f.2.4）に示したような母数間の対応をもつ一般模型（4f.1.2）の1つの特定の場合を得た．共分散構造（4f.2.4）をもつ観測値の模型（4f.2.3）は分散成分模型として知られている．分散成分の他の定式化は**4j**で与える．

理論的には，有限の対象物の場合における非復元抽出のようなもので（または）等確率または非等確率または α と β との間に相関があるような他の型の標本抽出を許してもよい．最も一般的な模型を考え，**4e.2** で与えたような2元データの分散分析によって，母数（すなわち共分散構造）を推定する方法を示そう．特定の場合における結果を導くためには，母数 $\sigma_e^2,\sigma_A^2,\sigma_B^2,\sigma_{AB}^2$ を，(α,β)についての特別な標本抽出機構にたいして(4f.2.4)にしたように，一般の母数 $\sigma^2,\rho_0,\rho_1,\rho_2,\rho_3$ であらわさねばならない．

4f.3　一般模型の処理

われわれの目的は，観測値にたいする一般的な共分散構造（4f.1.2）のもとで，分散分析表4e.2αにおける各項の期待値を決定することである．このために，変量 y_{ijk} から無相関の変量の組への直交変換をつくる．これは3段階でなされる．

第1段階　級における c 個の変量の各組に直交変換 0_1 をほどこし，新しい変量

$$u_{ij} = \frac{y_{ij}}{\sqrt{c}}, \qquad g_{ij2}, \cdots, g_{ijc} \tag{4f. 3. 1}$$

$$i = 1, \cdots, q, \quad j = 1, \cdots, p$$

を求める（ただし $y_{ij} = \sum_k y_{ijk}$）.

$$E(g_{ijk}) = 0, \quad V(g_{ijk}) = (1 - \rho_0)\sigma^2, \quad E(u_{ij}) = \sqrt{c}\,\mu_{ij}$$

$$V(u_{ij}) = (1 + \overline{c-1}\rho_0)\sigma^2, \qquad \mathrm{cov}(u_{ij}, u_{im}) = c\rho_1\sigma^2$$

$$\mathrm{cov}(u_{ij}, u_{mj}) = c\rho_2\sigma^2, \qquad \mathrm{cov}(u_{ij}, u_{mr}) = c\rho_3\sigma^2$$

であることを確かめよ.

第2段階　第 i 行の p 個の変量 $u_{i1}, \cdots, u_{ip}$ に直交変換 0_2 をほどこして新しい変量

$$z_{i1} = \sum \frac{u_{ij}}{\sqrt{p}}, \qquad z_{i2}, \cdots, z_{ip} \tag{4f. 3. 2}$$

$$i = 1, \cdots, q$$

を求める.（4f. 3. 1）のときと同じように，次のことが容易に確かめられる.

$$V(z_{i1}) = (1 + \overline{c-1}\rho_0 + c(p-1)\rho_1)\sigma^2$$

$$V(z_{ij}) = (1 + \overline{c-1}\rho_0 - c\rho_1)\sigma^2, \qquad j = 2, \cdots, p$$

$$\mathrm{cov}(z_{i1}, z_{m1}) = c(\rho_2 + \overline{p-1}\rho_3)\sigma^2$$

$$\mathrm{cov}(z_{ij}, z_{mj}) = c(\rho_2 - \rho_3)\sigma^2, \qquad j \neq 1$$

$$\mathrm{cov}(z_{ij}, z_{mr}) = 0, \qquad j \neq 1, \ r \neq 1.$$

第3段階　この複雑さを整然とするためには，第3そして最後の直交変換 0_3 が，q 個の変量 $z_{1j}, \cdots, z_{qj}$ の各組について必要とされ，新しい変量

$$v_{1j} = \sum \frac{z_{ij}}{\sqrt{q}}, \qquad v_{2j}, \cdots, v_{qj} \tag{4f. 3. 3}$$

$$j = 1, \cdots, p$$

が求められる．これらはすべて無相関であるが，次のように分散は等しくはない：

$$\sigma_{00} = V(v_{11}) = [1 + \overline{c-1}\rho_0 + c(p-1)\rho_1 + c(q-1)\rho_2 + c(p-1)(q-1)\rho_3]\sigma^2$$

$$\sigma_{10} = V(v_{i1}) = [1 + \overline{c-1}\rho_0 - c\rho_2 + c(p-1)(\rho_1 - \rho_3)]\sigma^2, \quad i > 1$$

$$\sigma_{01} = V(v_{1j}) = [1 + \overline{c-1}\rho_0 - c\rho_1 + c(q-1)(\rho_2 - \rho_3)]\sigma^2, \quad j > 1$$

$$\sigma_{11} = V(v_{ij}) = [1 + \overline{c-1}\rho_0 - c\rho_1 - c\rho_2 + c\rho_3]\sigma^2, \quad i, j \geqslant 2$$

$$\sigma_e^2 = V(g_{ijk}) = (1 - \rho_0)\sigma^2, \qquad k \geqslant 2. \tag{4f. 3. 4}$$

分散成分模型にたいしては，（4f. 2. 4）から，ρ_i, σ^2 に $\sigma_e^2, \sigma_A^2, \sigma_B^2, \sigma_{AB}^2$ を代入すると

$$\sigma_{00} = \sigma_e^2 + c\sigma_{AB}^2 + cp\sigma_A^2 + cq\sigma_B^2$$

$$\sigma_{10} = \sigma_e^2 + c\sigma_{AB}^2 + cp\sigma_A^2, \qquad \sigma_{01} = \sigma_e^2 + c\sigma_{AB}^2 + cq\sigma_B^2$$

$$\sigma_{11} = \sigma_e^2 + c\sigma_{AB}^2, \qquad \sigma_e^2 = \sigma_e^2 \tag{4f. 3. 5}$$

が得られる.

表 4f. 3α　一般模型（4f. 1. 2）のもとでの2元分散分析における平均平方の期待値

	D. F.	S. S.（もとの変量）	S. S.（新しい変量）	M. S.	M. S. の期待値
A	$q-1$	$\sum \frac{y_{i\cdot}^2}{pc} - \frac{y_{\cdot\cdot}^2}{pqc} = S_A$	$\sum_2^q v_{i1}^2$	V_A	$\sigma_{10} + c\delta_A^2$
B	$p-1$	$\sum \frac{y_{\cdot j}^2}{qc} - \frac{y_{\cdot\cdot}^2}{pqc} = S_B$	$\sum_2^p v_{1j}^2$	V_B	$\sigma_{01} + c\delta_B^2$
AB	$(p-1)(q-1)$	$* \quad = S_{AB}$	$\sum_2^p \sum_2^q v_{ij}^2$	V_{AB}	$\sigma_{11} + c\delta_{AB}^2$
級間	$pq-1$	$\sum\sum \frac{y_{ij}^2}{c} - \frac{y_{\cdot\cdot}^2}{pqc}$	$\sum_1^p \sum_1^q v_{ij}^2 - v_{11}^2$	V	$\sigma^2 + c\delta^2$
級内	$pq(c-1)$	$* \quad = S_e$	$\sum_2^c \sum_1^p \sum_1^q g_{ijk}^2$	V_e	σ_e^2

さて分散分析表 4f. 3α の平方和を新しい無相関の変量 v_{ij}, g_{ijk} であらわそう．最初の変換の結果として

$$\sum\sum\sum y_{ijk}^2 = \sum\sum \frac{y_{ij}^2}{c} + \sum\sum\sum g_{ijk}^2,$$

すなわち自由度 $pq(c-1)$ の級内の平方和は

$$S_e = \left(\sum\sum\sum y_{ijk}^2 - \frac{y_{\cdot\cdot}^2}{pqc}\right) - \left(\sum\sum \frac{y_{ij}^2}{c} - \frac{y_{\cdot\cdot}^2}{pqc}\right) = \sum\sum\sum g_{ijk}^2$$

であり，それゆえ $E(S_e) = E[pq(c-1)V_e] = \sum\sum\sum E(g_{ijk}^2) = pq(c-1)\sigma_e^2$.

〔$y_{\cdot\cdot} = \sum\sum\sum y_{ijk}$ であり，これは 4e. 3 では $y_{\cdot\cdot}$ となっている.〕

y_{ij} から v_{ij} への変換は直交である．それゆえ

$$\sum\sum v_{ij}^2 - v_{11}^2 = \sum\sum \frac{y_{ij}^2}{c} - \frac{y_{\cdot\cdot}^2}{pqc}$$

$$= \sum \frac{y_{i\cdot}^2}{pc} - \frac{y_{\cdot\cdot}^2}{pqc} + \sum \frac{y_{\cdot j}^2}{qc} - \frac{y_{\cdot\cdot}^2}{pqc} + c\sum\sum(\bar{y}_{ij} - \bar{y}_{i\cdot} - \bar{y}_{\cdot j} + \bar{y}_{\cdot\cdot})^2$$

$$= S_A + S_B + S_{AB}. \tag{4f. 3. 6}$$

$0_2, 0_3$ は直交変換であるから，

$$\sum_1^q \frac{y_{i\cdot}^2}{pc} - \frac{y_{\cdot\cdot}^2}{pqc} = \sum_2^q v_{i1}^2, \qquad \sum_1^p \frac{y_{\cdot j}^2}{qc} - \frac{y_{\cdot\cdot}^2}{pqc} = \sum_2^p v_{1j}^2,$$

したがって

$$S_{AB} = \sum_2^p \sum_2^q v_{ij}^2.$$

平均値 μ_{ij} についての恒等式（4f. 3. 6）を用いると

$$c(pq-1)\delta^2 = c\Big(\sum\sum \mu_{ij}^2 - \frac{\mu_{..}^2}{pq}\Big)$$

$$= c\Big(\sum \frac{\mu_{i.}^2}{p} - \frac{\mu_{..}^2}{pq}\Big) + c\Big(\sum \frac{\mu_{.j}^2}{q} - \frac{\mu_{..}^2}{pq}\Big)$$

$$+ c\sum\sum(\mu_{ij} - \bar{\mu}_{i.} - \bar{\mu}_{.j} + \bar{\mu}_{..})^2 \qquad (4f.3.7)$$

$$= c(q-1)\delta_A^2 + c(p-1)\delta_B^2 + c(p-1)(q-1)\delta_{AB}^2$$

が得られる．したがって

$$E[S_A] = E\Big(\sum_2^q v^2{}_1\Big) = \sum_2^q [V(v_{i1}) + \overline{E(v_{i1})}^2] \doteqdot (q-1)(\sigma_{10} + c\delta_A^2).$$

同様にして

$$E[S_B] = (p-1)(\sigma_{01} + c\delta_B^2), \qquad E[S_{AB}] = (p-1)(q-1)(\sigma_{11} + c\delta_{AB}^2).$$

これらの結果は分散分析表 4f.3α にまとめられている．ただし σ_{ij} と δ_i^2 はそれぞれ (4f.3.4) と (4f.3.7) で定義されたものである．

分散成分模型では，$\delta_A^2, \delta_B^2, \delta_{AB}^2$ はゼロであり，σ_{ij} は (4f.3.5) で与えられた値をもち，次の推定値が得られることを注意する．

$$\hat{\sigma}_e^2 = V_e, \qquad \hat{\sigma}_{AB}^2 = \frac{V_{AB} - V_e}{c}$$

$$\hat{\sigma}_A^2 = \frac{V_A - V_{AB}}{pc}, \qquad \hat{\sigma}_B^2 = \frac{V_B - V_{AB}}{qc}. \qquad (4f.3.8)$$

$\delta_A^2, \delta_B^2, \delta_{AB}^2$ がゼロではない一般の場合には，すべての未知母数を推定することはできない．

さて変量 y_{ijk} は多次元正規分布をすると仮定しよう．この場合には変換された変量 v_{ij}, g_{ijk} は独立で正規分布をする．したがって A, B, AB および級内による平方和 S_A, S_B, S_{AB}, S_e は，変量 v_{ij}, g_{ijk} の排他的な集合に依存しており，大抵の場合，すべて独立に分布する．かくして，

$$\frac{S_e}{\sigma_e^2} \sim \chi^2[pq(c-1)], \quad \text{中心 } \chi^2$$

$$\frac{S_A}{\sigma_{10}} \sim \chi^2\Big(q-1, \frac{c(q-1)\delta_A^2}{\sigma_{10}}\Big), \quad \delta_A^2 \neq 0 \text{ ならば非心 } \chi^2$$

$$\frac{S_B}{\sigma_{01}} \sim \chi^2\Big(p-1, \frac{c(p-1)\delta_B^2}{\sigma_{01}}\Big), \quad \delta_B^2 \neq 0 \text{ ならば非心 } \chi^2$$

$$\frac{S_{AB}}{\sigma_{11}} \sim \chi^2\Big((p-1)(q-1), \frac{c(p-1)(q-1)\delta_{AB}^2}{\sigma_{11}}\Big), \quad \delta_{AB}^2 \neq 0 \text{ ならば非心 } \chi^2.$$

これらの分布は未知母数に関するあらゆる推測問題の基礎となる．

たとえば分散成分模型では，$(S_A/\sigma_{10}) \sim \chi^2(q-1)$，$(S_e/\sigma_e^2) \sim \chi^2[pq(c-1)]$ であるから

$$\frac{V_A}{V_e}\frac{\sigma_e^2}{\sigma_{10}} \sim F[q-1,\ pq(c-1)] \qquad (4f.3.9)$$

であり，したがって比 σ_{10}/σ_e^2 の下方信頼限界は，(4f.3.9) の左辺を F のパーセント点と等しいとおく，つまり

$$\frac{V_A}{V_e}\frac{\sigma_e^2}{\sigma_{10}} = F_\alpha$$

とおくことにより求められる．

4g　統計的回帰の理論と応用

4g.1　回帰の概念（一般理論）

回帰の理論は，他の測定値または補助変量 $(x_1, \cdots, x_p) = \mathbf{x}'$ によって与えられる情報をもとにして，1つまたはそれ以上の変量 $(y_1, \cdots, y_q)$ の予測と関係がある．前者を独立または予測子変量，後者を従属または基準変量とよぶのが普通である．

予測は種々の実際の場面において必要である．気象学者は，1時点でとられた適当な大気測定値をもとにして，数時間先の天気予報をしたい．教育団体は適当な考査により特定の分野における候補者の能力を評価しなければならない．企業は，適当な検査装置を用いて，割り当てたい仕事にたいして能力のある人を募集したい．作物または動物の養育者は，子孫が最大の報酬または“利得”をもたらすような個体を選ぶことに興味がある．

これらすべての場面において，基準は将来において観測可能な変量であり，それらは決定をするため，適当な測定値によって予測されるべきものである．予測子をどのように選ぶべきか？

最小平均2乗誤差予測子　ただ1つの基準変量 y とベクトル $\mathbf{x}$ であらわされる p 個の予測子変量 $x_1, \cdots, x_p$ を考えよう．$f(x_1, \cdots, x_p)$ を y の予測子とする．

（i）<u>$M(\mathbf{x})$ を $E(y|\mathbf{x})$ とする．そうすると $E[y - f(\mathbf{x})]^2$ は $f(\mathbf{x}) = M(\mathbf{x})$ のとき最小である．</u>

条件つき期待値の定義から

$$E[(y-M)(M-f)] = E[(M-f)E(\overline{y-M}|\mathbf{x})] = 0$$

それゆえ

$$E(y-f)^2 = E(y-M+M-f)^2$$

$$= E(y-M)^2 + E(M-f)^2 \geqq E(y-M)^2. \qquad (4g.1.1)$$

$E(y-f)^2$ の下限は $f = M$ のとき到達される．したがって平均2乗誤差 m. s. e. (mean square error) を最小にする最良の予測子は，$\mathbf{x}$ が与えられたときの y の条件つき期待値 $M(\mathbf{x})$ であり，それは $x_1, \cdots, x_p$ にたいする y の回帰と呼ばれる．m. s. e. $E(y-M)^2$ は

$\mathbf{x}$ が与えられたときの y の条件つき分散の平均であり，$\sigma^2_{y\cdot\mathbf{x}}$ であらわす．

基準と最大の相関をもつ予測子　y と f との間の相関 $\rho(f,y)=\mathrm{cov}(y,f)/\sigma_f\sigma_y$ を最大にする f を見い出そう．

（ii）$M(\mathbf{x})$ は（i）におけるものであるとする．そうすると $\rho(y,M)$ は非負であり，任意の関数 f にたいして $\rho(y,M)\geqslant|\rho(y,f)|$.

任意の f にたいして

$$\mathrm{cov}(y,f)=E\{[f-E(f)]E(y|\mathbf{x})\}=E\{[f-E(f)]M\}=\mathrm{cov}(f,M) \tag{4g. 1. 2}$$

$f=M$ のときには，(4g. 1. 2) を用いると $\mathrm{cov}(y,M)=\mathrm{cov}(M,M)$．したがって $\rho(y,M)=\sigma_M/\sigma_y\geqslant 0$，ただし $\sigma^2_M=\mathrm{cov}(M,M)=V(M)$，$\sigma^2_y=V(y)$．さて

$$\rho^2(f,y)=\frac{[\mathrm{cov}(y,f)]^2}{\sigma^2_f\sigma^2_y}=\frac{[\mathrm{cov}(f,M)]^2}{\sigma^2_f\sigma^2_M}\cdot\frac{\sigma^2_M}{\sigma^2_y}$$
$$=\rho^2(f,M)\cdot\rho^2(y,M)\leqslant\rho^2(y,M). \tag{4g. 1. 3}$$

$\rho(f,y)$ の上限は $\rho(f,M)=1$ のとき到達され，そのときの f は M の線形関数である．またも回帰関数が答である．到達される最大の相関の2乗 σ^2_M/σ^2_y は**相関比**として知られており，$\eta^2_{y\mathbf{x}}$ であらわされる．定義より $0\leqslant\eta^2_{y\mathbf{x}}\leqslant 1$．次の分解（または分散分析）

$$E[y-E(y)]^2=E(y-M)^2+E[M-E(M)]^2$$
$$\sigma^2_y=\sigma^2_{y\cdot\mathbf{x}}\qquad+\sigma^2_M \tag{4g. 1. 4}$$

が得られる．それゆえ $\eta^2_{y\mathbf{x}}=1-(\sigma^2_{y\cdot\mathbf{x}}/\sigma^2_y)$ であり，これは，予測の誤差 $\sigma^2_{y\cdot\mathbf{x}}$ が0に近づくとき1に近づき，$\mathbf{x}$ の利用による誤差の減少がないとき零に近づく．かくして，もっと明確には $\eta^2_{y(x_1,\cdots,x_p)}$ または $\eta^2_{y(1,\cdots,p)}$ と書くべき $\eta^2_{y\mathbf{x}}$ は，y と $x_1,\cdots,x_p$ の間の関係または予測の正確さの尺度を与え，異なった場面を比較するときまたは与えられた問題において予測子変量の異なった選定の比較をするとき有用であろう．

選抜において期待成果を最大にすること　母集団の与えられた割合 α を残し，$\mathbf{x}$ の残りの分布における y の期待値を最大にするように，$\mathbf{x}$ の分布を断ち切りたいものとする．つまり，$\mathbf{x}$ をもとにして母集団の α 部分を，y の平均値が可能な最大をもつように特徴づけたい．もし $F(\mathbf{x},y)$ が同時分布関数ならば，

$$\iint_{w\times R}dF(\mathbf{x},y)=\alpha\qquad\text{のもとで}$$
$$\iint_{w\times R}y\,dF(\mathbf{x},y)\qquad\text{が最大である}$$

ような $\mathbf{x}$ の領域 w が必要である．ここで R は y の全範囲である．y に関して積分すると，問題は

$$\int_w dF(\mathbf{x})=\alpha\qquad\text{のもとで} \tag{4g. 1. 6}$$
$$\int_w M(\mathbf{x})dF(\mathbf{x})\qquad\text{を最大にすること} \tag{4g. 1. 5}$$

に帰着される．

（iii）（Cochran (1951)）　$M(\mathbf{x})\geqslant\lambda$ によって定義される領域 w は (4g. 1. 5) を最大にする．ただし λ は (4g. 1. 6) を満足するように選ばれる．

この結果は **7a. 2** で証明されるネイマン＝ピアソンの補助定理からすぐに得られる．またも答は回帰関数に基づいている．

線形回帰　回帰関数が $\mathbf{x}$ に関して線形である特別な場合は広く研究されている．$E(y-M)^2$ は M が回帰関数であるとき最小であるから，任意の線形関数 $\alpha+\beta_1x_1+\cdots+\beta_px_p$ を考え，

$$E(y-\alpha-\beta_1x_1-\cdots-\beta_px_p)^2=b^2+\boldsymbol{\beta}'\mathbf{C}\boldsymbol{\beta}-2\boldsymbol{\beta}'\boldsymbol{\sigma}_0+V(y) \tag{4g. 1. 7}$$

を最小にするように係数を決定することができる．ここで $\mathbf{C}$ は $x_1,\cdots,x_p$ の分散共分散行列，$\boldsymbol{\sigma}_0$ は $x_1,\cdots,x_p$ と y の共分散のベクトル，$b=\alpha-E(y)+\beta_1E(x_1)+\cdots+\beta_pE(x_p)$ である．b と $\boldsymbol{\beta}$ の最適な選定は

$$b^*=0,\qquad\mathbf{C}\boldsymbol{\beta}^*=\boldsymbol{\sigma}_0 \tag{4g. 1. 8}$$

である．というのは，$\boldsymbol{\beta}=\boldsymbol{\beta}^*+\boldsymbol{\delta}$ とおくと

$$b^2+\boldsymbol{\beta}'\mathbf{C}\boldsymbol{\beta}-2\boldsymbol{\beta}'\boldsymbol{\sigma}_0=(\boldsymbol{\beta}^*)'\mathbf{C}\boldsymbol{\beta}^*-2\boldsymbol{\sigma}_0'\boldsymbol{\beta}^*+\boldsymbol{\delta}'\mathbf{C}\boldsymbol{\delta}+b^2$$
$$\geqslant(\boldsymbol{\beta}^*)'\mathbf{C}\boldsymbol{\beta}^*-2\boldsymbol{\sigma}_0'\boldsymbol{\beta}^*+(b^*)^2$$

であることが容易に示されるからである．α の最適な値は

$$\alpha^*=E(y)-(\boldsymbol{\beta}^*)'E(\mathbf{x}) \tag{4g. 1. 9}$$

である．(4g. 1. 7) の最小値は

$$V(y)+(\boldsymbol{\beta}^*)'\mathbf{C}\boldsymbol{\beta}^*-2(\boldsymbol{\beta}^*)'\boldsymbol{\sigma}_0=\sigma_{00}-\boldsymbol{\sigma}_0'\mathbf{C}^{-1}\boldsymbol{\sigma}_0 \tag{4g. 1. 10}$$

である．ただし $\sigma_{00}=V(y)$．かくして，もし回帰が線形であるならば，変量の平均値，分散，共分散だけにより完全に規定され，変量の正確な分布の知識は必要でないことがわかる．

線形回帰の場合における相関比 $\eta^2_{y\mathbf{x}}$ は

$$\rho^2_{y\mathbf{x}}=1-\frac{\sigma^2_{y\cdot\mathbf{x}}}{\sigma^2_y}=\frac{\boldsymbol{\sigma}_0'\mathbf{C}^{-1}\boldsymbol{\sigma}_0}{\sigma_{00}} \tag{4g. 1. 11}$$

となり，変量の分散と共分散だけを用いて上のように定義される係数 ρ は，回帰が線形であってもなくても，**重相関係数**という特別な名前で知られている．もし線形回帰が成立しないならば，η^2 と (4g. 1. 11) に定義される ρ^2 は異なり，事実 $\eta^2\geqslant\rho^2$ である．

最良線形予測子　回帰が線形であるときには，最良の予測子は1次，2次の積率だけ

で決定されるということを知った．さて，回帰が線形であってもなくても，$\mathbf{x}$ について線形な y の予測子のうち最良なものは何であるかを調べよう．つまり，$\mathbf{x}$ のすべての関数の組を考えるかわりに，探究の範囲を $\mathbf{x}$ の線形関数だけに制限する．

(iv) (4g. 1. 8) と (4g. 1. 9) で決定される線形関数 $\alpha^* + \mathbf{x}'\beta^*$ は y の最小 m. s. e. 線形予測子である．それはまた，y と最大の相関をもつ線形関数である．

最初の部分は，事実，α^* と β^* は $E(y-\alpha-\beta'\mathbf{x})^2$ を最小にするようにきめられているということから明らか．

第2の部分を証明するために，

$$\mathrm{cov}(y, \beta'\mathbf{x}) = \beta'\sigma_0 = \beta'\mathbf{C}\beta^*$$
$$\mathrm{cov}(y, \mathbf{x}'\beta^*) = (\beta^*)'\mathbf{C}\beta^* = V(\mathbf{x}'\beta^*)$$

ということに注目しよう．したがって，C-S 不等式を用いると $(\beta'\mathbf{C}\beta^*)^2 \leqslant (\beta^{*\prime}\mathbf{C}\beta^*)(\beta'\mathbf{C}\beta)$ であることと，

$$\sigma_y\rho(y, \mathbf{x}'\beta^*) = \frac{\mathrm{cov}(y, \mathbf{x}'\beta^*)}{\sqrt{(\beta^*)'\mathbf{C}\beta^*}} = \sqrt{(\beta^*)'\mathbf{C}\beta^*}$$

ということに注目すると

$$\sigma_y^2\rho^2(y, \beta'\mathbf{x}) = \frac{(\beta'\mathbf{C}\beta^*)^2}{\beta'\mathbf{C}\beta} \leqslant (\beta^*)'\mathbf{C}\beta^* = \sigma_y^2\rho^2(y, \mathbf{x}'\beta^*)$$

がわかる．最大相関の2乗は

$$\rho_{y\mathbf{x}}^2 = \frac{\beta^{*\prime}\mathbf{C}\beta^*}{\sigma_{00}} = \frac{\sigma_0'\mathbf{C}^{-1}\sigma_0}{\sigma_{00}}$$

であり，これは前に導いた表現 (4g. 1. 11) である．$\eta_{y\mathbf{x}}$ は y と $\mathbf{x}$ の任意の関数との間の最大相関であり，$\rho_{y\mathbf{x}}$ は y と $\mathbf{x}$ の線形関数との間の最大相関であるから，$\eta_{y\mathbf{x}}^2 \geqslant \rho_{y\mathbf{x}}^2$ がわかる．係数 $\rho_{y\mathbf{x}}^2$ はもっと明確には $\rho_{y(x_1,\cdots,x_p)}^2$ または $\rho_{0(1\cdots p)}^2$ と書かれる．

ρ^2 の他の表現　(ρ_{ij}), $i, j = 0, 1, \cdots, p$ は変量 $y, x_1, \cdots, x_p$ の相関行列であるとし，(ρ^{ij}) は (ρ_{ij}) の逆行列であるとする．同様にして，(σ^{ij}) は，全部の変量の分散共分散行列 (σ_{ij}) の逆行列であるとする．方程式 $\mathbf{C}\beta^* = \sigma_0$ を詳細に書くと

$$\beta_1^*\sigma_{1i} + \cdots + \beta_p^*\sigma_{pi} = \sigma_{0i}, \qquad i = 1, \cdots, p \tag{4g. 1. 12}$$

が得られる．しかし定義から

$$\sigma^{10}\sigma_{1i} + \cdots + \sigma^{p0}\sigma_{pi} = -\sigma_{0i}\sigma^{00}, \qquad i = 1, \cdots, p \tag{4g. 1. 13}$$

が得られる．(4g. 1. 12) と (4g. 1. 13) を比較することにより

$$\beta_i^* = -\frac{\sigma^{i0}}{\sigma^{00}} = -\frac{\sqrt{\sigma_{00}}}{\sqrt{\sigma_{ii}}}\frac{\rho^{i0}}{\rho^{00}}$$

が得られる．さらに

$$\rho_{0(1\cdots p)}^2 = \frac{\sigma_0'\mathbf{C}^{-1}\sigma_0}{\sigma_{00}} = \frac{\sigma_0'\beta^*}{\sigma_{00}}$$

$$= -\frac{\sigma_{10}\sigma^{10} + \cdots + \sigma_{p0}\sigma^{p0}}{\sigma_{00}\sigma^{00}} = \frac{\sigma_{00}\sigma^{00} - 1}{\sigma_{00}\sigma^{00}}$$
$$= 1 - \frac{1}{\sigma_{00}\sigma^{00}} \tag{4g. 1. 14}$$
$$= 1 - \frac{1}{\rho^{00}} \qquad (\rho^{ij} \text{ によって表現された } \beta^* \text{ を代入することにより})$$

最後の式は $\rho_{0(1\cdots p)}^2$ が相関だけの関数であることを示している．

4g. 2　関連性の増加の尺度

実際問題として，余分の予測子変量を導入することにより基準変量 y の予測の精度が改善される程度を調べることが必要である．変量 $x_1, \cdots, x_k$ を用いたときの基準変量の予測の m.s.e.は $\sigma_y^2[1 - \eta_{0(1\cdots k)}^2]$ であり，$x_1, \cdots, x_p$, $p > k$ を用いたときは $\sigma_y^2[1 - \eta_{0(1\cdots p)}^2]$ である．後者は前者よりも大きくはなく，平均2乗誤差の減少分は

$$\sigma_y^2[\eta_{0(1\cdots p)}^2 - \eta_{0(1\cdots k)}^2]$$

であり，($x_1, \cdots, x_k$ に加うるに) 余分の変量 $x_{k+1}, \cdots, x_p$ の利用による減少の割合は

$$\frac{\eta_{0(1\cdots p)}^2 - \eta_{0(1\cdots k)}^2}{1 - \eta_{0(1\cdots k)}^2} \tag{4g. 2. 1}$$

となり，これは，偏相関比と呼ばれ $\eta_{0(k+1\cdots p)\cdot(1\cdots k)}^2$ であらわす．これは，y と $x_{k+1}, \cdots, x_p$ の直接的な関係の測度である相関比 $\eta_{0(k+1\cdots p)}^2$ と区別されるべきである．

もし線形予測子だけを考えるならば，m. s. e. の減少の割合の対応する表現は

$$\rho_{0(k+1\cdots p)\cdot(1\cdots k)}^2 = \frac{\rho_{0(1\cdots p)}^2 - \rho_{0(1\cdots k)}^2}{1 - \rho_{0(1\cdots k)}^2} \tag{4g. 2. 2}$$

であり，これは，いわば，$x_1, \cdots, x_k$ による関係を除去することによる $x_{k+1}, \cdots, x_p$ にたいする y の偏重相関 (の2乗) とよんでよい．

偏関連性のこの尺度は，少しちがった方法ででも考えられる．$x_1, \cdots, x_k$ に基づいて y を予測するときの誤差は $e = y - M(x_1, \cdots, x_k)$ である．ただし M は真の回帰 (または線形予測子) をあらわす．もしこの誤差が，(いま用いている変量 $x_1, \cdots, x_k$ に加うるに) 余分の変量 $x_{k+1}, \cdots, x_p$ を考えることによりある点までは予測可能であるとするならば，後者をつけ加えることは有用であろう．そうするとその尺度は，誤差 e の予測における m. s. e.，つまり $x_1, \cdots, x_p$ にたいする e の相関比 (または線形予測の場合には重相関) に関係する．e の予測子は

$$E(e|x_1, \cdots, x_p) = M(x_1, \cdots, x_p) - M(x_1, \cdots, x_k) \tag{4g. 2. 3}$$

であり，これの分散は

$$\sigma_y^2[\eta_{0(1\cdots p)}^2 - \eta_{0(1\cdots k)}^2] \tag{4g. 2. 4}$$

であるが，e の全分散は

$$\sigma_y^2[1-\eta_{0(1\cdots k)}^2] \tag{4g. 2. 5}$$

であるということを確かめることは教育的である．それゆえ要求される相関比は (4g. 2. 5) にたいする (4g. 2. 4) の比であり，これは (4g. 2. 1) で定義された尺度と同じである．η を ρ でおきかえることにより，線形予測子にたいする対応する表現が得られる．y と，$x_1, \cdots, x_{p-1}$ を除去した x_p との偏関連性を測る係数 $\rho_{0(p)\cdot(1\cdots p-1)}^2$ には特別な興味がある．というのは，これによって，現存する組に加えられる個々の変量の重要性を判断することができるからである．(4g. 2. 2) から，$1-\rho_{0(1\cdots k)}^2$ にたいして公式 (4g. 1. 14) を用いることにより

$$1-\rho_{0(p)\cdot(1\cdots p-1)}^2=\frac{1-\rho_{0(1\cdots p)}^2}{1-\rho_{0(1\cdots p-1)}^2}=\frac{\begin{vmatrix}1 & \cdots & \rho_{1\,p-1}\\ & \cdots\cdots\cdots & \\ \rho_{1\,p-1} & \cdots & 1\end{vmatrix}}{\rho^{00}|\rho_{ij}|_{p-1}} \tag{4g. 2. 6}$$

が得られる．ただし $|\rho_{ij}|_{p-1}$ は $y, x_1, \cdots, x_{p-1}$ だけの相関の行列式をあらわす．さらに，逆行列の要素の性質から

$$\rho^{00}\rho^{pp}-(\rho^{0p})^2=\frac{\begin{vmatrix}1 & \cdots & \rho_{1\,p-1}\\ & \cdots\cdots\cdots & \\ \rho_{1\,p-1} & \cdots & 1\end{vmatrix}}{|\rho_{ij}|_p} \tag{4g. 2. 7}$$

が成立するから，公式 (4g. 2. 6) より

$$\rho_{0(p)\cdot(1\cdots p-1)}^2=\frac{(\rho^{0p})^2}{\rho^{00}\rho^{pp}}$$

が得られ，この表現は，偏関連性の尺度は記号 0 と p に関して対称であり，$\rho_{0(p)\cdot(1\cdots p-1)}^2$ と $\rho_{p(0)\cdot(1\cdots p-1)}^2$ とは同じ値をもつことを示している．平方根をとり，2番目の記号の括弧をとって

$$\rho_{0p\cdot(1\cdots p-1)}=\frac{\rho^{0p}}{\sqrt{\rho^{00}\rho^{pp}}} \tag{4g. 2. 8}$$

と書き，係数 (4g. 2. 8) に，y と，$x_1, \cdots, x_{p-1}$ を除去した x_p との間の偏相関という特別な呼び名を与えている．

偏相関 $\rho_{0p\cdot(1\cdots p-1)}$ は，y と x_p から $x_1, \cdots, x_{p-1}$ にたいするそれぞれの回帰式を引いたあとの残差

$$\left.\begin{aligned}e_1&=y-(\alpha+\beta_1x_1+\cdots+\beta_{p-1}x_{p-1})\\ e_2&=x_p-(\alpha'+\beta_1'x_1+\cdots+\beta_{p-1}'x_{p-1})\end{aligned}\right\} \tag{4g. 2. 9}$$

の間の相関として一般に定義されている．こうして定まる相関は確かに (4g. 2. 8) であるということを示すのは良い演習となるであろう．しかし (4g. 2. 8) の定義を導く接近のほうがより直観的，自然的，一般的であるように思える．

4g. 3 の例で，変量の分布の正規性の仮定のもとで，重相関，偏相関にたいする有意性検定を考える．

4g. 3　頭蓋容量の予測（1つの実例）

問題は頭蓋骨の容量を推定することであるが，その頭蓋骨はこわれていたりまたはいたんでいたりして，その容量を直接測定することはできないものとする．このような場合，少なくともいくつかの外形の測定値があるならば，ある精度でもって容量を予測することができる (Rao and Shaw (1948e))．

必要となる基本データは，大事に保存されたたくさんの頭蓋骨について，外形の測定値に加うるに容量の完全な記録であり，これらから，予測のための回帰関数を推定する．

回帰式の選定　頭蓋容量 C を予測するための3つの重要な測定値は眉間－後頭部の長さ L，頭頂骨最大幅 B，底後頭部－前頂の高さ H' である．体積を推定するためには，次の型の回帰式を仮定するのが適当である．

$$C=\gamma L^{\beta_1}B^{\beta_2}H'^{\beta_3} \tag{4g. 3. 1}$$

ここで $\gamma, \beta_1, \beta_2, \beta_3$ は求めるべき定数である．これらの変数を

$$y=\log_{10}C,\quad x_1=\log_{10}L,\quad x_2=\log_{10}B,\quad x_3=\log_{10}H'$$

と変換することにより，(4g. 3. 1) を

$$y=\alpha+\beta_1x_1+\beta_2x_2+\beta_3x_3 \tag{4g. 3. 2}$$

と書くことができる．ただし $\alpha=\log_{10}\gamma$．幸いにして，(4g. 3. 1) を線形式 (4g. 3. 2) に容易に帰着させることができた．なぜ，L, B, H' にたいして型

$$C=\alpha+\beta_1L+\beta_2B+\beta_3H' \tag{4g. 3. 3}$$

の線形回帰関数を採用しなかったのかという疑問が起るであろう．これに答えるためには，容量の予測における2つの関数 (4g. 3. 1) と (4g. 3. 3) の性能を調べなければならない．われわれはある意味において近い予測を与えるものを選ぶ．式の誤った選定ははなはだしく精度の悪い予測を導くので，われわれは経験を用い，合理的な式に到達する前にたくさんの式について試さねばならない．

定数の推定　(4g. 3. 2) のように，式が未知の定数に関して線形であるときには，未知定数を推定するのに最小2乗法が用いられる．(4g. 3. 2) を

$$\alpha'+\beta_1(x_1-\bar{x}_1)+\cdots+\beta_3(x_3-\bar{x}_3)$$

の形に書き直す．ここで $\bar{x}_1, \bar{x}_2, \bar{x}_3$ は測定 x_1, x_2, x_3 にたいする観測平均であり

$$\alpha'=\alpha+\beta_1\bar{x}_1+\beta_2\bar{x}_2+\beta_3\bar{x}_3$$

である．

測定値の組 $(y_r, x_{1r}, x_{2r}, x_{3r})$ が n 組あるとしよう．

$$\sum_{r=1}^{n}[y_r-\alpha'-\beta_1(x_{1r}-\bar{x}_1)-\beta_2(x_{2r}-\bar{x}_2)-\beta_3(x_{3r}-\bar{x}_3)]^2$$

を最小にすることによって得られる正規方程式は，$\alpha', \beta_1, \beta_2, \beta_3$ の推定値を a', b_1, b_2, b_3 と書くと

$$na'=\sum y_r$$
$$b_1S_{11}+b_2S_{12}+b_3S_{13}=S_{01}$$
$$b_1S_{12}+b_2S_{22}+b_3S_{23}=S_{02}$$
$$b_1S_{13}+b_2S_{23}+b_3S_{33}=S_{03}$$

であることを確かめよ．ただし S_{ij} は普通の修正された積和であって

$$S_{ij}=\sum_r(x_{ir}-\bar{x}_i)(x_{jr}-\bar{x}_j)=\sum x_{ir}x_{jr}-n\bar{x}_i\bar{x}_j$$
$$S_{0i}=\sum_r(y_r-\bar{y})(x_{ir}-\bar{x}_i)=\sum y_rx_{ir}-n\bar{y}\bar{x}_i.$$

Hooke (1926) によって得られたファリングドン街族からの $n=86$ の男の頭蓋骨についての測定値を用いて，次の値が計算された．

$$\bar{y}=3.1685,\quad \bar{x}_1=2.2752,\quad \bar{x}_2=2.1523,\quad \bar{x}_3=2.1128$$

$$(S_{ij})=\begin{pmatrix}0.01875 & 0.00848 & 0.00684\\ 0.00848 & 0.02904 & 0.00878\\ 0.00684 & 0.00878 & 0.02886\end{pmatrix}$$

$$S_{01}=0.03030,\quad S_{02}=0.04410,\quad S_{03}=0.03629$$

$$S_{00}=\sum y_r-n\bar{y}^2=0.12692$$

これらはすべて以後の解析に必要となる準備計算である．〔原著では $S_{00}=0.02781$ となっているが，以後の計算では S_{00} は 0.12692 として使われているので，後者を採用した．〕

行列 (S_{ij}) の逆行列を (C_{ij})［記号 (S^{ij}) を使うべきであるが，統計の文献では，これは，要素 (C_{ij}) をもつ C 行列としてすでに知られている］であらわすと

$$(C_{ij})=\begin{pmatrix}64.21 & -15.57 & -10.49\\ -15.57 & 41.71 & -9.00\\ -10.49 & -9.00 & 39.88\end{pmatrix}$$

である．母数の推定数（正規方程式の解）は，(C_{ij}) の最初の行を用いて

$$b_1=64.21S_{01}-15.57S_{02}-10.49S_{03}=0.878,$$

(C_{ij}) の2番目の行を用いて

$$b_2=-15.57S_{01}+41.71S_{02}-9.00S_{03}=1.041$$

である．同様にして $b_3=0.733$，最後に $a'=\bar{y}=3.1685$，かくして α の推定値は

$$a=\bar{y}-b_1\bar{x}_1-b_2\bar{x}_2-b_3\bar{x}_3=-2.618$$

として与えられる．それゆえ，もとの測定値を用いての推定式は

$$C=0.00241L^{0.878}B^{1.041}H'^{0.733} \tag{4g. 3. 4}$$

である．［ファリングドン街の頭蓋骨の容量は，からしの種をきつく詰め込み，それらを秤量することによりきめられた．(4g. 3. 4) は厳密には，この方法できめられる容量を予測する場合にのみ適用可能である．］

残差平方和は

$$R_0^2=S_{00}-\sum b_iS_{0i}=0.12692-0.09911=0.02781$$

である．予測の精度または従属変数と独立変数との間の関連性の尺度を与える重相関は

$$R^2=\frac{b_1S_{01}+\cdots+b_3S_{03}}{S_{00}}=\frac{0.09911}{0.12692}=0.7809$$

によって推定される．x_1, x_2, x_3 が与えられたときの y の分散 σ^2 は

$$s^2=\frac{R_0^2}{n-p-1}=\frac{0.02781}{82}=0.0003391$$

によって推定される．ここで p は独立変数の数である．b_i の分散，共分散は公式

$$V(b_i)=C_{ii}\sigma^2,\qquad \mathrm{cov}(b_i,b_j)=C_{ij}\sigma^2$$

によって計算される．これで最小2乗法による未知母数，標準誤差の推定を終る．

仮説の検定　回帰式の推定や使用において起こるいくつかの問題を考えよう．これらはすでに求めている推定値を用いて述べられ，また調べられる．

1.　従属変数と独立変数の間には有意な関連性があるか？　母数でいうと，仮説は $\beta_1=\beta_2=\beta_3=0$ である．これはまた，真の重相関 (4g. 1. 11) が零であるという陳述と同値である．

分散分析の手法を用いる．自由度 $n-p-1=82$ の条件をつけない平方和 R_0^2 は，0.02781 としてすでに求めている．もし仮説 $\beta_1=\beta_2=\beta_3=0$ が真であるならば，$\sum(y_i-\alpha)^2$ の最小値 R_1^2 は $\sum y_i^2-n\bar{y}^2=S_{00}=0.12692$ であり，自由度 $n-1$ をもつ．残差平方和の減少 $R_1^2-R_0^2$ は回帰によるものであり，自由度 $p=3$ をもつ．平方和の分析は表 4g. 3α に示される．表 4g. 3α の分散比は，重相関の2乗 R^2 の値，標本の大きさ n，独立変数の数 p から，R^2 の計算値を用い，公式

$$F=\frac{R^2}{1-R^2}\cdot\frac{n-p-1}{p}=\frac{0.7809}{1-0.7809}\cdot\frac{86-3-1}{3}=97.41$$

によって簡単に求められることがわかる．それゆえ考えている仮説の検定を，観測された

表 4g. 3α　仮説 $\beta_1=\beta_2=\beta_3=0$ の検定　(原典：Rao and Shaw (1948e))

	D. F.	S. S.	M. S.	F
回　帰	3	0.09911	0.033037	97.41
残　差	82	0.02781	0.0003391	
合　計	85	0.12692		

重相関の値だけを知って行うことができる．自由度3，82の分散比97.41は1％水準で有意であり，これは予測式における変数の有用性を示す．

2.　3つの寸法が予測式において同じ次数で現われるかどうかを調べてみよう．推定値から頭骨最大幅の指数 b_2 は他のものよりも高いことがわかる．これは，幅の一定比の増加は，長さや高さの対応する増加よりも容量にたいして大きい価値をもつということを意味する．

この問題の検討に関連する仮説は

$$\beta_1 = \beta_2 = \beta_3 = \beta$$

である．$\sum[y_i - \alpha - \beta(x_{1i} + x_{2i} + x_{3i})]^2$ の最小値を求めなければならない．β の推定値を与える正規方程式は

$$[S_{11} + S_{22} + S_{33} + 2(S_{12} + S_{23} + S_{31})]b = Q = (S_{01} + S_{02} + S_{03})$$

$$0.12485b = 0.11069, \qquad b = 0.8866$$

である．自由度 $n-2$ をもつ最小値は

$$(\sum y_i^2 - n\bar{y}^2) - bQ = 0.12692 - 0.09814 = 0.02878$$

である．表4g.3β の F 比は小さいから，$\beta_1, \beta_2, \beta_3$ が異なっていると結論する根拠はない．その差は，たとえあっても，小さいようであり，これについて何かはっきりしたことを言う前に多数の測定値を集めることが必要であろう．進化論者は，幅は頭蓋骨の他の大きさよりも相対的に増加が大きいと信じている．もしこれが本当であるならば，頭蓋容量が幅にどれぐらい影響されるかを調べることは興味がある．

予測の問題に関する限り，$\beta_1 = \beta_2 = \beta_3$ を仮定して得られる式

$$C = 0.002342(LBH')^{0.8866}$$

は，この仮定なしで導かれた式と同じぐらい有用であろう．β の推定値 b の分散の推定値は $\sigma'^2/\sum\sum S_{ij}$ である．ただし σ'^2 は，表4g.3β に与えられた自由度84の平方和に基づく推定値である．

表 4g.3β　仮説 $\beta_1 = \beta_2 = \beta_3$ の検定

	D. F.	S. S.	M. S.	F
一様性からの偏差	2	0.00097	0.000485	1.430
残　差	82	0.02781	0.0003391	
合　計	84	0.02878		

3.　型 $C = \alpha' LBH'$ の単純な式が頭蓋容量を予測するのに時々用いられる．このような式の妥当性の検定は，仮説

$$\beta_1 = \beta_2 = \beta_3 = 1$$

を検定することと同値である．

$\beta_1 = \beta_2 = \beta_3 = 1$ であるときの $\sum(y_i - \alpha - \beta_1 x_{1i} - \beta_2 x_{2i} - \beta_3 x_{3i})^2$ の最小値は

$$(\sum y_i^2 - n\bar{y}^2) + S_{11} + S_{22} + S_{33} + 2(S_{12} + S_{23} + S_{31})$$
$$- 2(S_{01} + S_{02} + S_{03}) = 0.03039$$

であり，自由度 $n-1$ をもつ．残差は自由度 $n-4$ をもつから，自由度3をもつ差が仮説からの偏差によるものである．

自由度3，82の比2.544は5％有意水準のほんのわずか下である．もっと適当な資料でこの問題を調べるのは興味あることであろう．

表4g.3γ において，仮説からの偏差による平方和は，

$$\sum\sum S_{ij}(b_i - 1)(b_j - 1) = 0.00258$$

のように，推定値 b_1, b_2, b_3 の期待値1, 1, 1からの偏差の複合した尺度を与える式から直接計算することもできる．この式の値は表4g.3γ に与えられたものと同じである．

表 4g.3γ　仮説 $\beta_1 = \beta_2 = \beta_3 = 1$ の検定

	D. F.	S. S.	M. S.	F
$\beta_1 = \beta_2 = \beta_3 = 1$ からの偏差	3	0.00258	0.00086	2.544
残　差	82	0.02781	0.0003391	
合　計	85	0.03039		

4.　β 係数がそれぞれ1とは異なるということが明らかになると，用いられる3つの寸法の指数は一様ではないが，指数の合計が3になるかどうかを調べることは少し興味がある．これは仮説 $\beta_1 + \beta_2 + \beta_3 = 3$ の検定を要求する．偏差の最良の推定量は $b_1 + b_2 + b_3 - 3 = 2.652 - 3 = -0.348$ である．

$$V(b_1 + b_2 + b_3 - 3) = (\sum\sum C_{ij})\sigma^2 = 75.68\sigma^2.$$

自由度1，82の比は

$$\frac{(0.348)^2}{75.68} \times \frac{1}{0.0003391} = 4.72$$

であり，これは5％水準で有意である．これは，予測式の次元の数は3ではないということを示す．

5.　余分の変数の導入が予測の精度を増加させるかどうかを検定したい場合がしばしばある．たとえば，前の例において L, B に加えて H' が必要かどうかを検定することができる．これは $\beta_3 = 0$ であるかどうかを検定することと同値である．推定値 $b_3 = 0.733$ は分散 $C_{33}\sigma^2$ をもつ．自由度1，82の比は

$$\frac{b_3^2}{C_{33}\hat{\sigma}^2} = \frac{(0.733)^2}{39.88} \times \frac{1}{0.0003391}$$

$$= \frac{0.01347}{0.0003391} = 39.72$$

である．ただし $\hat{\sigma}^2$ にたいしては自由度82に基づく推定値が用いられる．これは1％水準で有意であり，H' もまた関係があることを示している．もし b_3 が有意でなかったとすると，b_3 による平方和

$$\frac{b_3^2}{C_{33}} = 0.01347$$

を残差平方和0.02781につけ加え，自由度83に基づく平方和0.04128を得ることができる．これは σ^2 の推定値として0.0004973を与える．

もし b_3 は零であると断定されるならば，式

$$y = \alpha + \beta_1 x_1 + \beta_2 x_2$$

から出発して，b_1, b_2 の最良の推定値は修正されねばならない．しかしながら，新たに出発する必要はない．C 行列

$$\begin{matrix} 64.21 & -15.57 & -10.49 \\ -15.57 & 41.71 & -9.00 \\ -10.49 & -9.00 & 39.88 \end{matrix}$$

を，最後の行から出発し，枢軸として C_{33} を用い，枢軸縮約の方法によって縮約する．かくして

$$\begin{bmatrix} 64.21 - \frac{(10.49)^2}{39.88} & -15.57 - \frac{(9.00)(10.49)}{39.88} \\ -15.57 - \frac{(9.00)(10.49)}{39.88} & 41.71 - \frac{(9.00)^2}{39.88} \end{bmatrix} = \begin{bmatrix} 61.45 & -17.94 \\ -17.94 & 39.68 \end{bmatrix}$$

が b_1, b_2 を評価するための縮約された C 行列を与える．

$$b_1 = 61.45S_{01} - 17.94S_{02} = 1.071$$

$$b_2 = -17.94S_{01} + 39.68S_{02} = 1.206$$

$$a = \bar{y} - b_1\bar{x}_1 - b_2\bar{x}_2 = -1.864$$

残差平方和は

$$\sum (y_i - \bar{y})^2 - b_1S_{01} - b_2S_{02} = 0.04128$$

である．これは，残差に b_3 による平方和を加えて得た値と一致しており，したがって b_1, b_2 の計算のチェックを与える．b_1, b_2 の分散・共分散行列はこの新しい C 行列の σ^2 倍である．

さらに変数を落すならば，枢軸縮約の方法をさらに行わねばならない．各段階での縮約された行列は残っている変数に対応する C 行列を与える．枢軸縮約の方法を行う前に，除去する変数の順番に C 行列を書いておくと混乱がさけられるであろう．かくして x_1, x_3, x_3, x_4 がもとの変数であり，x_2 と x_4 を除去したいとするならば，C 行列を

$$\begin{matrix} C_{22} & C_{24} & C_{21} & C_{23} \\ C_{42} & C_{44} & C_{41} & C_{43} \\ C_{12} & C_{14} & C_{11} & C_{13} \\ C_{32} & C_{34} & C_{31} & C_{33} \end{matrix}$$

のように書くとよい．これは2番目，4番目の行と列を最初の2つの位置にもってくることにより得られる．そうするとこの行列は，前むき枢軸縮約の方法によって2段階で縮約することができる．

6.　<u>個々の場合における値の信頼区間</u>　x_1, x_2, x_3 を個々の場合における独立変数の観測値とし，y を予測されるべき値とする．$Y(x_1, x_2, x_3)$ を x_1, x_2, x_3 の与えられた値にたいする推定回帰式の値とし，$y - Y(x_1, x_2, x_3)$ を考えよう．

$$V(y - Y) = \sigma^2 + V(Y)$$

$$V(y) = V[\bar{y} + b_1(x_1 - \bar{x}_1) + \cdots + b_3(x_3 - \bar{x}_3)]$$

$$= \sigma^2\left\{\frac{1}{n} + \sum\sum (x_i - \bar{x}_i)(x_j - \bar{x}_j)C_{ij}\right\}$$

したがって

$$t = (y - Y) \div s\left\{1 + \frac{1}{n} + \sum\sum (x_i - \bar{x}_i)(x_j - \bar{x}_j)C_{ij}\right\}^{\frac{1}{2}}$$

は，σ の推定値 s に対応して，自由度 $n - p - 1$ の t 分布をする．t に，t の上側，下側の $\alpha/2$ 確率点を代入すると，これに対応して y の2つの値が得られ，これが未知の値にたいする $1 - \alpha$ 信頼区間を与える．$L = 198.5$, $B = 147$, $H' = 131$ をもつ頭蓋骨，すなわち $x_1 = 2.298$, $x_2 = 2.167$, $x_3 = 2.117$ は，回帰の値

$$Y = \bar{y} + b_1(x_1 - \bar{x}_1) + \cdots + b_3(x_3 - \bar{x}_3) = 3.2069$$

$$s^2\left\{1 + \frac{1}{n} + \sum\sum (x_i - \bar{x}_i)(x_j - \bar{x}_j)C_{ij}\right\} = 0.0003391(1.04187)$$

$$= 0.0003533$$

をもつ．0.95信頼区間を求めるためには，方程式は

$$\frac{y - 3.2069}{(0.0003533)^{\frac{1}{2}}} = \pm 2.00$$

である．ここで2.00は自由度82の $|t|$ の近似的な5％値である．または，この個体の頭蓋骨の対数容量の0.95信頼区間は

$$3.2069 \pm 2.00(0.01880) = (3.1693, 3.2445)$$

である．y の信頼区間から，y の1対1関数の信頼区間が求められるというのは1つの重要な性質である．したがって容量（$= \log y$ ではない）の0.95信頼区間は（1476, 1755）である．この区間はやや広く，任意特定の場合における予測はよくないであろうということ

とを示している．

7.　従属変数（対数容量）の平均の推定　　問題が，前に選ばれた L, B, H' をもつ頭蓋骨の平均対数容量の推定であるならば，推定値は

$$Y = \bar{y} + b_1(x_1 - \bar{x}_1) + \cdots + b_3(x_3 - \bar{x}_3) = 3.2069$$

として得られ，それの分散は

$$V(Y) = \sigma^2\left\{\frac{1}{n} + \sum\sum(x_i - \bar{x}_i)(x_j - \bar{x}_j)C_{ij}\right\}$$

$$= 0.0003391(0.04187) = 0.00001420 \quad (\sigma \text{ にたいし } s \text{ を用いた})$$

である．0.95 信頼区間は

$$3.2069 \pm 2.00(0.0000142)^{\frac{1}{2}} = (3.1994, 3.2144)$$

である．ここで2.00は自由度82の $|t|$ の5％値である．残念ながら，平均容量の信頼区間を対数容量の平均の信頼区間から求めることはできない．もしも容量それ自身が補助測定値に関して線形回帰をもつならばこの困難は起らない．

推定値 Y の分散にたいする前の公式において，推定値の精度は，x_1, x_2, x_3 が予測式を推定するのに使われた標本の平均 $\bar{x}_1, \bar{x}_2, \bar{x}_3$ にどれぐらい近いかに関係することがわかる．事実，分散は，新しい標本についての測定値が前の標本の平均値と一致するときの推定値にたいして最小（σ^2/n に等しい）である．予測の精度は x_1, x_2, x_3 が平均値から離れるにつれて減少し，x_1, x_2, x_3 が標本において観測された値の範囲の外に落ちるならば，予測は完全に信頼できないものになるであろう．

ただ1つの変数を用いての予測では，回帰式とその分散は

$$Y = a + b(x - \bar{x})$$

$$V(Y) = \left\{\frac{1}{n} + \frac{(x - \bar{x})^2}{S_{11}}\right\}\sigma^2$$

である．ただし S_{11} は観測された標本における x の修正された平方和である．前と同じように，予測の精度は平均の近くで高い．この公式は標本における x の散布状態をあらわす S_{11} にも関係しており，散布状態が大きければ大きいほど，任意の x にたいする推定値の精度は高い．それゆえ，予測式をつくるための標本を選ぶに際しては，最良の予測式をつくろうと思うならば x の範囲の両端のところで観測をするべきである．これは，疑いもなく，回帰式が線形であるということがはっきりとわかっているときには有効に実行することができる理論的に妥当な政策である．このような方法で集められたデータは，回帰が線形であるか否かを判定するのは適当ではなく，回帰の線形性が一般的であると信じる理由はない．生物測定の経験では，回帰は非常に近似的に線形であり，線形性からの偏差は大標本においてのみ検出可能である．もしそうであるならば，データにたいして適合された回帰直線は真の回帰関数への1つの近似にすぎず，理想的な曲線に最もよく適合する直線

が得られるようにデータはとられるべきである．これのための最良の計画は，全範囲のなかから x を選ぶこと，またはむしろ x の異なった値がそれぞれの確率であらわれ，それらが適合する直線の決定に影響を及ぼすように x を確率的に選ぶことである．

母集団の平均容量の推定　　上に求めた予測式は，L, B, H' だけの測定値がとれる頭蓋骨の標本から平均頭蓋容量を推定するために用いることもできる．この目的のためには2つの方法がある．おのおのの頭蓋骨の頭蓋容量を推定しこれらの推定値の平均を計算することと，標本の L, B, H' の平均値に直接予測式を適用することである．これら2つの方法が同じ結果を与えるかどうかを知ることは興味がある．この目的のために，L, B, H' の測定値はあるが C のそれはないようなファリングドン街族のほかの29個の男の頭蓋骨にたいして平均頭蓋容量の推定値をつくった．

これら29個の頭蓋骨にたいして L, B, H' の平均はそれぞれ 191.1，143.1，129.0 である．これらの平均値にたいして公式 $C = 0.00241L^{0.878}B^{1.041}H'^{0.733}$ を適用し，C の平均は 1498.3 であると推定する．29個の頭蓋骨個々にたいして C を推定し29個の推定値の平均をとると，C の平均値は 1498.2 に等しいという推定値が得られる．

同様な推定値はモールフィールズ族（Hooke (1926)）の22個の男の頭蓋骨にたいして計算された．ただしこの場合には4つ全部の測定値があった．これら22個の頭蓋骨にたいして，L, B, H' の平均はそれぞれ 189.5，142.5，128.8 であり，C の平均の推定値として 1479.0 を与える．22個の頭蓋骨個々にたいして C を推定し平均を計算するならば，C の平均の推定値として 1480.0 を得る．かくして2つの方法はほぼ同じ推定値を与えるようである．

小さい頭蓋骨だけが保存されるのか？　　次のことはいくらか興味がある．86個の測定された値から計算したファリングドン街族の C の平均値は 1481.3 であるが，L, B, H' の測定値はあるが C のそれはないような29個の頭蓋骨から公式によって推定された C の平均値は 1498.3 である．

また，モールフィールズ族にたいしては，44個の測定値に基づく L の平均は 189.2 であり，46個の測定値に基づく B の平均は 143.0 であり，34個の測定値に基づく H' の平均は 129.8 である．これらの平均値にたいして公式を適用すると（前に示したようにある信頼度でもってそうしてよいから），C の平均の推定値として 1490.7 を得る．上手に保存された頭蓋骨の22個の測定値から計算した C の平均はわずかに 1473.8 である．

これらの結果は，頭蓋容量が測定できない程度にこわれている頭蓋骨は，もとのままで残っているものよりも概して大きいということを示唆する．

これは重要な争点を提起する．頭蓋容量の発表された平均値はまずい低目の推定値ではないか？　これらの値を修正するための適当な方法はないか？　1つの方法は，予測式を

つくるためにだけ C, L, B, H' についての観測値のある標本を用いることであろう．前に考察したように，用いられる回帰関数の性質が適当であるならば，予測式は，標本は母集団から確率的に抽出されてないけれども，測定項目すべてについて観測値のある標本から求めることができた．たとえば，もし小さい頭蓋骨だけが保存されるならば，得られる測定値は頭蓋骨の母集団から確率的にとられたとは厳密にはいえない．このような材料は関係式を定めるためにだけ用いられる．こうして予測式が得られたならば，平均容量の推定値を得るために，現存する測定値全部から計算された L, B, H' の平均値を代入する．この値は，現存する頭蓋容量の測定値の平均よりも高いが，L, B, H' だけの測定値をもつ頭蓋骨の平均に基づく予測値よりも小さいであろう．

低目の推定の程度は大きい頭蓋骨の崩壊の割合に関係する．これは族ごとにちがうであろう．したがって平均容量の正しい比較のためには前に指摘した修正を適用するべきであろう．

4g. 4　回帰式の一様性の検定

2つの族からつくられた回帰関数が同じであるかどうかを検定することが非常にしばしば必要である．したがって，異なった族からの頭蓋容量の予測式が

$$\alpha' + \beta_1' x_1 + \beta_2' x_2 + \beta_3' x_3$$

であるならば，次の2つの型の仮説が検定される．

（i）　$\alpha = \alpha'$，$\beta_1 = \beta_1'$，$\beta_2 = \beta_2'$，$\beta_3 = \beta_3'$

（ii）　α が α' に等しいかどうかとは無関係に，$\beta_1 = \beta_1'$，$\beta_2 = \beta_2'$，$\beta_3 = \beta_3'$.

(i) が正しいならば，回帰関数は両方の族で全く同じである．(ii) が正しいならば，回帰関数は，定数のちがいは別として同じである．1つの族からつくられた予測式をまったく異なった族からの標本にたいして使わねばならない場合がたくさん起こるから，これら2つの仮説は適切である．これの利用の極端かつやや野心的な例は，近代の人の身長とその人の大腿骨の長さとを結びつける公式を用いて，化石の大腿骨の長さから大昔の人の身長の予測をすることである (K. Pearson (1898))．このような手順のある種の正当性は，最初の仮説が類似の場合において正しいことが示されるならば，得られるであろう．まず，両方の族にたいして予測式が推定されているときの検定手順を取り扱う．

第2の族にたいして得られた数値は次のようであったとする．

標本の大きさ：　n'

平均値：　$\bar{x}_1', \bar{x}_2', \bar{x}_3', \bar{y}'$

修正された積和：

$$\sum (x_{ir}' - \bar{x}_i')(x_{jr}' - \bar{x}_j') = S_{ij}'$$

$$\sum (x_{ir}' - \bar{x}_i')(y_r' - \bar{y}') = S_{0i}'$$

$$\sum (y_r' - \bar{y}')^2 = S_{00}'$$

これらは回帰関数

$$y = a' + b_1' x_1 + b_2' x_2 + b_3' x_3$$

を決定するのに十分である．残差平方和は

$$R_0^2 = S_{00}' - b_1' S_{01}' - b_2' S_{02}' - b_3' S_{03}' \quad \text{（第2の標本にたいして）}$$
$$+ S_{00} - b_1 S_{01} - b_2 S_{02} - b_3 S_{03} \quad \text{（第1の標本にたいして）}$$

であり，自由度 $(n' - 4) + (n - 4) = (n + n' - 8)$ をもつ．ここで4は推定された定数の数をあらわす．

さて，2つの標本を一緒にし，それらを大きさ $n + n'$ のただ1つの標本と考え，回帰直線と残差平方和を決める．必要な数値はすでにある数値から計算することができる．

標本の大きさ：　$n + n'$

平均値：

$$\frac{(n\bar{x}_i + n'\bar{x}_i')}{(n + n')} = \bar{x}_i''$$

修正された積和：

$$S_{ij}'' = S_{ij} + S_{ij}' + \frac{nn'}{n + n'}(\bar{x}_i - \bar{x}_i')(\bar{x}_j - \bar{x}_j')$$

$$S_{0i}'' = S_{0i} + S_{0i}' + \frac{nn'}{n + n'}(\bar{x}_i - \bar{x}_i')(\bar{y} - \bar{y}')$$

$$S_{00}'' = S_{00} + S_{00}' + \frac{nn'}{n + n'}(\bar{y} - \bar{y}')^2$$

$b_1'', b_2'', \cdots$ を回帰係数とすると，自由度 $n + n' - 4$ の残差平方和 R_1^2 は

$$S_{00}'' - b_1'' S_{01}'' - b_2'' S_{02}'' - b_3'' S_{03}''$$

である．表 4g. 4α のような分散分析をする．

表 4g. 4α　回帰係数の一様性を検定するための分散分析

変動因	D. F.	S. S.
仮説からの偏差	4	*（引き算による）
個々の回帰（残差）	$n + n' - 8$	R_0^2
共通の回帰（残差）	$n + n' - 4$	R_1^2

仮説からの偏差による平均平方の個々の回帰による残差にたいする比の有意性を検定する．

目的が b 係数だけの一様性を検定することであれば，次の量

$$S_{0i}''' = S_{0i} + S_{0i}', \qquad S_{ij}''' = S_{ij} + S_{ij}'$$

を計算し，方程式

$$S_{0i}''' = b_1''' S_{1i}''' + b_2''' S_{2i}''' + b_3''' S_{3i}''', \qquad i = 1, 2, 3$$

から定数 b_1''', b_2''', b_3''' を求め，自由度 $n + n' - 5$ の残差平方和 R_2^2 を求める．ここで5は，回帰係数の数と各族の異なった定数の数2を加えたものをあらわす．

$$R_2^2 = \sum (y_r - \bar{y})^2 + \sum (y_r' - \bar{y}')^2 - b_1''' S_{01}''' - b_2''' S_{02}''' - b_3 S_{03}'''.$$

検定は，自由度3，$n + n' - 8$ の分散比

$$\frac{R_2^2 - R_0^2}{3} \div \frac{R_0^2}{n + n' - 8}$$

に関係する．ここで3は比較される回帰係数の数をあらわす．

任意個数の予測変数の場合の解析や任意個数の回帰関数の比較への拡張はほとんど明らかである．

生物学上のデータでは，測定値の相互の相関や変動性はすべての類似の族にたいして近似的に同じであるということがしばしば見られる．この場合には，回帰式における係数 $\beta_1, \beta_2, \beta_3$ はあまりちがわないであろう．一方，平均値は族ごとに多少ちがい，この場合には定数項は同じではないであろう．これは $\beta_1 = \beta_1', \beta_2 = \beta_2', \cdots$ が与えられたとき $\alpha = \alpha'$ であるかどうかの異なった問題を提起する．これのための検定は前に計算した平方和から直ちに得られる．適当な統計量は，自由度1，$n + n' - 5$ の分散比

$$\frac{R_1^2 - R_2^2}{1} \div \frac{R_2^2}{n + n' - 5}$$

である．前の仮説が正しいならば，y の平均値の差は他の変数 x_1, x_2, x_3 の差によって完全に説明されることになる．十分な数の測定値を考えていると，すでに考えている変数の組と独立な他の測定値によって寄与される余分の差は無視できるほど小さいようである．このような場合には，両方の族の分散共分散行列の一様性は，全体として回帰関数の一様性を保証するのに十分である．1つまたは2つの変数に基づく予測式をこのように用いるときには細心の注意が必要である．しかしながら，大昔の人の身長を予測するというカール・ピアソンによってくわだてられたこのような統計的冒険は，われわれが彼の論文 (K. Pearson (1898)) の最後の文章に同意するならば正当と認められる．

"科学的研究は最終のものではない．それは単に，著者の責任において，データからひき出される最も確からしい結論をあらわすにすぎない．もっと広い範囲の事実やもっと洗練された解析・実験・観測は新しい公式や新しい理論を導くであろう．これが科学的進歩の本質である."

4g. 5　指定された回帰関数にたいする検定

4g. 3では，対数容量にたいする回帰関数は長さ，幅，高さの対数に関して線形である

と仮定された．もし，少なくとも独立変数のある与えられた値の組にたいして，従属変数の観測値が2つ以上あるならば，このような仮定の妥当性を検定することができる．

表4g. 5α はインド各地に住む人々の鼻示数の平均値，各地の気温と相対湿度の年平均を与えたものである．鼻示数の修正された総平方和は11, 140. 209で，自由度は344-1であることがわかっている．

表 4g. 5α　住民の鼻示数とその地域の気温と湿度

地域	標本の大きさ (n)	鼻示数 平均 ($\bar{y}$)	鼻示数 合計	気温 年平均 ($\bar{t}$)	気温 合計 ($n\bar{t}$)	相対湿度 年平均 ($\bar{h}$)	相対湿度 合計 ($n\bar{h}$)
Assam	36	83. 0	2, 988. 0	72. 6	2, 613. 6	85	3, 060
Orissa	40	80. 4	3, 216. 0	80. 3	3, 212. 0	69	2, 760
Bihar	30	80. 1	2, 403. 0	74. 8	2, 244. 0	88	2, 640
Malabar	45	77. 0	3, 465. 0	80. 2	3, 609. 0	81	3, 645
Bombay	26	76. 2	1, 981. 2	77. 6	2, 017. 6	66	1, 716
Madras	35	75. 9	2, 656. 5	81. 8	2, 863. 0	76	2, 660
Punjab	28	71. 4	1, 999. 2	76. 4	2, 139. 2	63	1, 764
United Province	32	80. 8	2, 585. 6	77. 2	2, 470. 4	69	2, 208
Andhra	41	76. 8	3, 148. 8	80. 3	3, 292. 3	69	2, 829
Ceylon	31	80. 3	2, 490. 0	80. 2	2, 486. 2	82	2, 542
合計	344		26, 933. 3		26, 947. 3		25, 824
平均		78. 294		78. 335		75. 070	

気温と相対湿度にたいする鼻示数の回帰について何の仮定もしないならば，群間，群内の分散分析は表4g. 5β のようになる．群間の平均平方は非常に大きく，鼻示数に確かな差があることを示している．これらの差を，形

$$y = \alpha + \beta_1 t + \beta_2 h$$

の，気温 t と相対湿度 h にたいする鼻示数 y の線形回帰式で説明することができるであろうか？　β_1, β_2 の推定値 b_1, b_2 を求めるための正規方程式は

表 4g. 5β　鼻示数にたいする分散分析

	D. F.	S. S.	M. S.
群間	9	$\sum (\bar{y})(n\bar{y}) - \frac{\sum n\bar{y}}{\sum n} \sum n\bar{y} = 3, 169. 900$	352. 21
群内	334	* (引き算による) 7, 970. 309	23. 863
合計	343	11, 140. 209	

$$S_{01} = b_1 S_{11} + b_2 S_{12}$$
$$S_{02} = {}_1S_{12} + b_2 S_{22}$$

である．ただし

$$S_{01} = \sum (n\bar{y}\bar{t}) - \frac{(\sum n\bar{t})(\sum n\bar{y})}{\sum n} = -1,072.40$$

$$S_{02} = \sum (n\bar{y}\bar{h}) - \frac{(\sum n\bar{h})(\sum n\bar{y})}{\sum n} = -4,586.57$$

$$S_{12} = \sum (n\bar{t}\bar{h}) - \frac{(\sum n\bar{t})(\sum n\bar{h})}{\sum n} = -2,334.91$$

$$S_{11} = \sum (n\bar{t}\bar{t}) - \frac{(\sum n\bar{t})(\sum n\bar{t})}{\sum n} = 2,721.06$$

$$S_{22} = \sum (n\bar{h}\bar{h}) - \frac{(\sum n\bar{h})(\sum n\bar{h})}{\sum n} = 22,042.32.$$

前の方程式を解くと，b_1, b_2 として

$$b_1 = -0.237113, \qquad b_2 = 0.182963$$

を得る．これらの値を用いて回帰分析を行うと表 4g. 5γ のようになる．回帰に関する仮説が正しいならば，表 4g. 5β の“群内”と表 4g. 5γ の“回帰のまわりの残差”とから得られる平均平方は同じ大きさであろう．有意な差は仮説を論駁することになる．

自由度 7，334 の比 12. 4 は 1 %水準で有意であるから，気温と相対湿度にたいする鼻示数の回帰は線形であるとは考えられない．また表 4g. 5γ から，自由度 2，341 の分散比 18. 55 は有意であることがわかる．しかし，これは，鼻示数が各人の住んでいる場所の気候条件に完全に関係するということを意味しない．観測された差は，気候の差で説明されるものよりもっと複雑であるであろう，または気候との関係の様子は非常に複雑なものであろう．

表 4g. 5γ　回　帰　分　析

	D. F.	S. S.	M. S	F
回帰による	2	$b_1S_{01} + b_2S_{02} = 1,093.45$	546. 72	18. 55
回帰のまわりの残差	341	* 10,046. 759 (引き算による)	29. 463	
合　計	343	11,140. 209		

父と娘の身長の分布の場合のように，ある場合には，1つの変数の他の変数にたいする回帰が線形であるかどうかを検定したい．このような検定にたいしては，独立変数の範囲を適当な数の級区間に分け，従属変数の分散を級間，級内で分析しなければいけない．級

内変動の推定値を得るためには，少なくともいくつかの級が2つ以上の観測値をもつことが必要である．回帰分析は級区間を用いないで行えばよい．または，もしデータがすでに群分けされているならば，級区間の中点を，その級における従属変数の各観測値にたいする独立変数の値とする．最終的な検定は表 4g. 5δ のようにしてなされる．

表 4g. 5δ　指定された回帰関数にたいする検定

	D. F.	S. S.	M. S.	F
指定された回帰からの偏差	7	2,076. 45*	296. 636	12. 4
群　内	334	7,970. 309	23. 863	
回帰のまわりの残差	341	10,046. 759		

* 引き算によって求める．

このほかの例題については Fisher (1925)，Linnik (1961)，Plackett (1960)，Scheffe (1959)，Williams (1959)，Yates (1949) の本を参照されたい．実験計画への応用は Box (1954)，Rao (1947c, 1956a, 1971B2) の論文や，Cochran and Cox (1957)，Kempthorne (1952) の本にある．

4g. 6　制約のある回帰

ある選択問題においては (Kempthorne and Nordskog, (1959) を参照)，選択された母集団において，基準変量 Y_1 については最大の利得をもち，他の基準変量 $Y_2, \cdots, Y_k$ とはそのままの状態を保つような補助（または予測子）変量 $X_1, \cdots, X_p$ に基づく線形選択指数が必要である．または，いいかえれば，$I = \mathbf{L'X} = L_1X_1 + \cdots + L_pX_p$ とするとき，条件 $\rho_{IY_2} = \cdots = \rho_{IY_k} = 0$ のもとで ρ_{IY_1} が最大であるような線形関数（指数）I を見いだすことが必要である．ただし ρ_{UV} は変量 U と V との相関をあらわす．

$\mathbf{C}_i$ を Y_i と $(X_1, \cdots, X_p)$ との共分散のベクトルとする．そうすると代数的問題は $\mathbf{L'C}_2 = \cdots = \mathbf{L'C}_k = 0$，$\mathbf{L'\Sigma L} = 1$ という条件のもとで $(\mathbf{L'C}_1)^2$ を最大にすることである．ここで Σ は $\mathbf{X}$ の分散共分散行列である．式

$$(\mathbf{L'C}_1)^2 - \lambda_1(\mathbf{L'\Sigma L} - 1) - 2\lambda_2\mathbf{L'C}_2 - \cdots - 2\lambda_k\mathbf{L'C}_k \qquad (4g.\ 6.\ 1)$$

を考えよう．ただし $\lambda_1, \cdots, \lambda_k$ はラグランジュ乗数である．$\mathbf{L}$ に関して微分をし零と等しいとおくと

$$\lambda_1\Sigma\mathbf{L} = \mu_1\mathbf{C}_1 - \lambda_2\mathbf{C}_2 - \cdots - \lambda_k\mathbf{C}_k \qquad (4g.\ 6.\ 2)$$

が得られる．ここで $\mathbf{L'C}_1$ を μ_1 と書いた．または最適の $\mathbf{L}$ について

$$\mathbf{L} \propto \Sigma^{-1}(\mathbf{C}_1 - \mu_2\mathbf{C}_2 - \cdots - \mu_k\mathbf{C}_k) \qquad (4g.\ 6.\ 3)$$

が得られる．条件 $\mathbf{L}'\mathbf{C}_i = 0$ より，$\mu_2, \cdots, \mu_k$ の満足するべき方程式

$$\mathbf{C}_i'\Sigma^{-1}\mathbf{C}_2\mu_2 + \cdots + \mathbf{C}_i'\Sigma^{-1}\mathbf{C}_k\mu_k = \mathbf{C}_i'\Sigma^{-1}\mathbf{C}_1 \qquad (4g.6.4)$$

$$i = 2, \cdots, k$$

が得られる．$\hat{\mu}_2, \cdots, \hat{\mu}_k$ を1つの解とすると，選択指数は

$$I = (\mathbf{C}_1' - \hat{\mu}_2\mathbf{C}_2' - \cdots - \hat{\mu}_k\mathbf{C}_k')\Sigma^{-1}\mathbf{X} \qquad (4g.6.5)$$

となり，もし必要ならば，定数をかけて分散が1になるように標準化をすればよい．

条件 $\mathbf{L}'\mathbf{C}_i \geqslant 0$，$i = 2, \cdots, k$，$\mathbf{L}'\Sigma\mathbf{L} = 1$ のもとで $\mathbf{L}'\mathbf{C}_1$ を最大にするもっと一般的な問題が Rao (1962d, 1964a) に解かれている．

4h　2組の母数をもつ最小2乗法の一般的問題

4h.1　補　助　変　数

異なった食物を与えた動物の群の成長率を比較したいものとする．処理をうける動物が，年齢，初期体重，血統のような成長率に影響を与える観測可能な性質に関してすべて同じであるときに限り，成長率の観測された差を食物のせいにすることができる．事実，もし異なった食物を与えられる動物の群がこれらの性質に関して異なっているならば，これらの差を除去した後，成長率を比較することが望ましい．

しかしながら，もしこれらの性質に関して異なっているかもわからない動物を異なった処理をうける群に確率的に割り当てるならば，実験には偏りがはいらないということを注意しておく．この操作は，同じ処理を受けた動物の成長率のちがいから計算される残差変動を大きくし，したがって実験の効率を下げる．

もしこれら付加的変数の大きさがわかっているならば，処理とは無関係にそれらに起因する差を，群内・群間の両方から除去し，処理の真の効果を高い効率で検定することができる．この手法に関係する計算技法は補助変動の調整として知られている．

残差変動の有意な減少にたいして，いま考えている効果が補助変数の影響をうけるということがわかっていなければいけない．これは，補助変数を記録するための費用や時間に相当の考慮をしなければいけないような実験研究の場合には重要であり，**4h.3** でとりあげる例で示すようにして検定することができる．

一方，選ばれた補助変数は考えている処理の影響をうけてはいけない．ときどき，異なった処理をうけた作物の収量の差を評価するのに，枝の数や葉の量のような補助変数が選ばれる．このような手順を使う際には相当の注意が必要である．

4h.2　共　分　散　分　析

補助変数 $c_1, \cdots, c_k$ にたいする y の回帰は線形であると仮定すると，考えている母数 $(\beta_1, \cdots, \beta_m) = \boldsymbol{\beta}'$ と回帰係数 $(\gamma_1, \cdots, \gamma_k) = \boldsymbol{\gamma}'$ を含む観測方程式は

$$E(y_i) = x_{i1}\beta_1 + \cdots + x_{im}\beta_m + \gamma_1 c_{1i} + \cdots + \gamma_k c_{ki}, \qquad i = 1, \cdots, n$$

または行列記号では

$$E(\mathbf{Y}) = \mathbf{X}\boldsymbol{\beta} + \mathbf{C}\boldsymbol{\gamma} \qquad (4h.2.1)$$

と書くことができる．ここで，$\mathbf{X}$ は前と同じように計画行列であり，$\mathbf{C}$ は，その i 列 $\mathbf{C}_i$ が i 番目の補助変数についての n 個の観測値をあらわす $n \times k$ の行列である．それゆえ，補助変数についての観測値を固定して考えるならば，模型 (4h.2.1) は，$m + k$ 個の未知母数 $(\boldsymbol{\beta}' \vdots \boldsymbol{\gamma}')$，新しい“計画行列”$(\mathbf{X} \vdots \mathbf{C})$，観測値ベクトル $\mathbf{Y}$ をもつ最小2乗法の理論の模型と同じである．

補助変数の調整においても，母数の推定や線形仮説の検定においても，新しい問題は起こらない．しかしながら，実際の場面において興味あるいくつかの仮説の検定を行うことができる計算手順を与えよう．観測方程式 (4h.2.1) に対応する正規方程式は

$$\left.\begin{aligned} \mathbf{X}'\mathbf{X}\boldsymbol{\beta} + \mathbf{X}'\mathbf{C}\boldsymbol{\gamma} &= \mathbf{X}'\mathbf{Y} \\ \mathbf{C}'\mathbf{X}\boldsymbol{\beta} + \mathbf{C}'\mathbf{C}\boldsymbol{\gamma} &= \mathbf{C}'\mathbf{Y} \end{aligned}\right\} \qquad (4h.2.2)$$

である．$\hat{\boldsymbol{\beta}}_0$ を，補助変数を無視した正規方程式

$$\mathbf{X}'\mathbf{X}\boldsymbol{\beta}_0 = \mathbf{X}'\mathbf{Y}$$

の1つの解とし，$\hat{\boldsymbol{\beta}}_i$ を，$\mathbf{Y}$ の代りに i 番目の補助変数の値を考えた（すなわち，$\mathbf{Y}$ のところに $\mathbf{C}_i$ を代入した）正規方程式

$$\mathbf{X}'\mathbf{X}\boldsymbol{\beta}_i = \mathbf{X}'\mathbf{C}_i$$

の1つの解とする．次のような残差平方和，残差積和を定義する．

$$\begin{aligned} R_{00} &= (\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}_0)'(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}_0) = \mathbf{Y}'\mathbf{Y} - \hat{\boldsymbol{\beta}}_0'\mathbf{X}'\mathbf{Y} \\ R_{ii} &= (\mathbf{C}_i - \mathbf{X}\hat{\boldsymbol{\beta}}_i)'(\mathbf{C}_i - \mathbf{X}\hat{\boldsymbol{\beta}}_i) = \mathbf{C}_i'\mathbf{C}_i - \hat{\boldsymbol{\beta}}_i'\mathbf{X}'\mathbf{C}_i, \qquad i = 1, 2, \cdots, k \\ R_{0i} &= (\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}_0)'(\mathbf{C}_i - \mathbf{X}\hat{\boldsymbol{\beta}}_i) = \mathbf{Y}'\mathbf{C}_i - \hat{\boldsymbol{\beta}}_0'\mathbf{X}'\mathbf{C}_i, \qquad i = 1, 2, \cdots, k \\ R_{ij} &= (\mathbf{C}_i - \mathbf{X}\hat{\boldsymbol{\beta}}_i)'(\mathbf{C}_j - \mathbf{X}\hat{\boldsymbol{\beta}}_j) = \mathbf{C}_i'\mathbf{C}_j - \hat{\boldsymbol{\beta}}_i'\mathbf{X}'\mathbf{C}_j, \qquad i, j = 1, \cdots, k \end{aligned} \qquad (4h.2.3)$$

(4h.2.2) の最初の方程式に $\hat{\boldsymbol{\beta}}_i'$ をかけ，(4h.2.2) の2番目の方程式の i 番目の成分から引くと，回帰母数だけの方程式

$$R_{1i}\gamma_1 + \cdots + R_{ki}\gamma_k = R_{0i}, \qquad i = 1, \cdots, k \qquad (4h.2.4)$$

を得る．$\hat{\boldsymbol{\gamma}}$ が (4h.2.4) の解であるならば，正規方程式 (4h.2.2) を満足する解 $\hat{\boldsymbol{\beta}}$ は

$$\hat{\boldsymbol{\beta}} = \hat{\boldsymbol{\beta}}_0 - \hat{\gamma}_1\hat{\boldsymbol{\beta}}_1 - \cdots - \hat{\gamma}_k\hat{\boldsymbol{\beta}}_k$$

であることがわかり，これは，補助変数のため $\hat{\boldsymbol{\beta}}_0$ になされるべき調整を与える．したがって模型 (4h.2.1) に適当な残差平方和は

$$R_0^2 = R_{00} - \hat{\gamma}_1 R_{01} - \cdots - \hat{\gamma}_k R_{0k} \qquad (4h.2.5)$$

であり，自由度は $n-r-k$ である．ただし $r = \mathrm{rank}\,\mathbf{X}$ ($|\mathbf{C}'\mathbf{C}| \neq 0$ を仮定している)．

仮説 $\gamma = 0$ の検定　もし $\gamma = 0$ ならば，補助変数にたいする調整は有利ではない．したがってこの仮説を検定することは重要である．

$$R_1^2 = \min_{\gamma=0,\beta} (\mathbf{Y} - \mathbf{X}\beta - \mathbf{C}\gamma)'(\mathbf{Y} - \mathbf{X}\beta - \mathbf{C}\gamma) = R_{00}$$

であることを認めると，仮説からの偏差による平方和は

$$R_1^2 - R_0^2 = \hat{\gamma}_1 R_{01} + \cdots + \hat{\gamma}_k R_{0k} \qquad (4h.2.6)$$

であり，自由度は k である．$\gamma = 0$ を検定するための分散比は

$$\frac{R_1^2 - R_0^2}{k} \div \frac{R_0^2}{n-r-k} \qquad (4h.2.7)$$

である．

仮説 $\mathbf{H}'\beta = \xi$ の検定

$$R_2^2 = \min_{\mathbf{H}'\beta=\xi} (\mathbf{Y} - \mathbf{X}\beta - \mathbf{C}\gamma)'(\mathbf{Y} - \mathbf{X}\beta - \mathbf{C}\gamma)$$

を求めなければいけない．β に制約 $\mathbf{H}'\beta = \xi$ があるとき，(4h.2.3) におけるような残差平方和，残差積和を $R'_{00}, R'_{0i}, R'_{ij}$ とする．方程式 (4h.2.4) において R'_{ij} を用いたときの解を $\hat{\gamma}'_1, \cdots, \hat{\gamma}'_k$ とする．そうすると

$$R_2^2 = R'_{00} - \hat{\gamma}'_1 R'_{01} - \cdots - \hat{\gamma}'_k R'_{0k}. \qquad (4h.2.8)$$

分散比検定は

$$\frac{R_2^2 - R_0^2}{s} \div \frac{R_0^2}{n-r-k} \qquad (4h.2.9)$$

である．ただし s は仮説 $\mathbf{H}'\beta = \xi$ の自由度であり，R_0^2 は，β または γ についての仮説がないときの純粋な残差平方和 (4h.2.5) である．これでひととおりの理論は終りである．4h.3 の問題においてこの方法をくわしく説明する．

4h.3　1つの例題

次のデータは，畜舎，性，与えられた食物の型で分類された30匹の豚の初期体重と成長率を示す．

問題は，初期体重を除去して食物の効果を研究することである．

解析の最初のステップは，従属，独立両方の変数の平方和，積和を分析することである．積和の分析は，平方和における1つの変数の平方を，積和に含まれている変数の積におきかえるという規則を採用して，平方和の分析と同じ方法でなされる．

総平方和と総積和は

表 4h.3α　解析のためのデータ (Wishart (1938))

畜舎	処理	性	初期体重 (w)	ポンドであらわした1週間あたりの成長率 (g)
I	A	G	48	9.94
	B	G	48	10.00
	C	G	48	9.75
	C	H	48	9.11
	B	H	39	8.51
	A	H	38	9.52
II	B	G	32	9.24
	C	G	28	8.66
	A	G	32	9.48
	C	H	37	8.50
	A	H	35	8.21
	B	H	38	9.95
III	C	G	33	7.63
	A	G	35	9.32
	B	G	41	9.34
	B	H	46	8.43
	C	H	42	8.90
	A	H	41	9.32
IV	C	G	50	10.37
	A	H	48	10.56
	B	G	46	9.68
	A	G	46	10.98
	B	H	40	8.86
	C	H	42	9.51
V	B	G	37	9.67
	A	G	32	8.82
	C	G	30	8.57
	B	H	40	9.20
	C	H	40	8.76
	A	H	43	10.42

$$\left.\begin{aligned} \sum g^2 - \bar{g}\sum g &= 16.6068 \\ \sum wg - \bar{w}\sum g &= 78.979 \\ \sum w^2 - \bar{w}\sum w &= 1108.70 \end{aligned}\right\} \text{自由度 } 29$$

である．w_i, g_i は i 番目の畜舎の6個の観測値の合計をあらわすとすると，畜舎の平方和，積和は

$$\left.\begin{aligned}\frac{1}{6}\sum g_i^2-\bar{g}\sum g &= 4.9607\\ \frac{1}{6}\sum g_iw_i-\bar{g}\sum w &= 40.324\\ \frac{1}{6}\sum w_i^2-\bar{w}\sum w &= 605.87\end{aligned}\right\}\text{自由度 4}$$

である．同様にして，食物，性の平方和と積和を求めることができる．

w_{ij}, g_{ij} は i 番目の型の食物の j 番目の性の5つの観測値の合計をあらわすとすると，食物と性の同時効果の平方和，積和は

$$\left.\begin{aligned}\frac{1}{5}\sum\sum g_{ij}^2-\bar{g}\sum g &= 3.2422\\ \frac{1}{5}\sum\sum g_{ij}w_{ij}-\bar{g}\sum w &= -0.885\\ \frac{1}{5}\sum\sum w_{ij}^2-\bar{w}\sum w &= 59.90\end{aligned}\right\}\text{自由度 5}$$

である．これらから食物（自由度2），性（自由度1）の対応する数値を引くと，食物×性 の交互作用の数値が得られる．表 4h. 3β は全体の分析を与える．誤差の行は合計から

表 4h. 3β　分散・共分散分析

	D. F.	g^2	wg	w^2
畜　舎	4	4.9607	40.324	605.87
食　物	2	2.3242	−0.171	5.40
性	1	0.4538	−4.813	32.03
食物×性	2	0.4642	4.099	22.47
誤　差	20	$8.4039 = R_{00}$	$39.540 = R_{01}$	$442.93 = R_{11}$
合　計	29	16.6068	78.979	1108.70

の引き算で求める．推定すべき回帰係数はただ1つである：

$$R_{11}\gamma = R_{01}$$

$$442.93\gamma = 39.540, \qquad \hat{\gamma} = 0.089269.$$

比は1％水準で有意であるから，補助変数の変動を除去することにより，もっと効率よく比較を行うことができる．

表 4h. 3γ　上の誤差の行から回帰の検定

	D. F.	S. S.	M. S.	F
回　帰	1	$\hat{\gamma}R_{01} = 3.5297$	3.5297	13.76
残　差	19	$R_{00}-\hat{\gamma}R_{01} = 4.8742$	0.2565	
合　計	20	$R_{00} = 8.4039$		

仮説が食物間に差がないということを規定するならば，残差平方和と残差積和は，食物

と誤差に対応する行を加えることによって得られる．

$$R'_{00} = 10.7281, \qquad R'_{01} = 39.369, \qquad R'_{11} = 448.33.$$

新しい回帰係数は，

$$\gamma' R'_{11} = R'_{01}, \qquad \hat{\gamma}' = 39.369/448.33 = 0.087813$$

である．仮説が正しいときの残差平方和は

$$R'_{00}-\hat{\gamma}'R'_{01} = 7.2710, \qquad \text{自由度は 21}$$

である．比は5％水準で有意である．補助変動の調整をしないで食物の検定をするために

表 4h. 3δ　初期体重の効果を除去したのち，食物の差の検定

	D. F.	S. S.	M. S.	F
食　物	2	* 2.3968	1.1984	4.67
		（引き算による）		
残　差	19	$R_{00}-\hat{\gamma}R_{01} = 4.8742$	0.2565	
食物＋誤差	21	$R'_{00}-\hat{\gamma}'R'_{01} = 7.2710$		

は，自由度2，20の比

$$F = \left(\frac{2.3242}{2}\right)\left(\frac{20}{8.4039}\right) = 2.76$$

をつくらなければいけなく，これは5％水準で有意ではない．Fを計算するために使われた数値は分散分析表（表 4h. 3β）からとられている．食物に起因する差は，補助変動が除去されたとき検出することができた．同様にして，性や交互作用のような他の効果も検定することができる．

4i　線形推定の統一理論

4i. 1　一般逆行列に関する基本的補助定理

模型 $(\mathbf{Y}, \mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{G})$ における $\boldsymbol{\beta}$ の推定問題は，$\mathbf{G}$ が正則であるという仮定のもとで 4a で考えた．しかしながら，その目的にたいして展開された最小2乗法の理論は $\mathbf{G}$ が特異であるときには使えない．

この節では，この問題にたいする2つの統一的接近を与える．それらは $\mathbf{G}$ が特異であろうとあるまいとすべての場合に適用可能である．1つは，不偏推定量の分散を最小にすることによって求める直接的接近である．もう1つは，最小2乗法の理論の類似である．

まず，統一理論において基本的役割を果たす分割行列の一般逆行列についての補助定理を証明する．

（i）　**基本的補助定理**　$\mathbf{G}$は次数nの n. n. d. 行列，$\mathbf{X}$は$n\times m$行列であるとする．さらに

$$\begin{pmatrix}\mathbf{G} & \mathbf{X}\\ \mathbf{X}' & \mathbf{0}\end{pmatrix}^{-}=\begin{pmatrix}\mathbf{C}_1 & \mathbf{C}_2\\ \mathbf{C}_3 & -\mathbf{C}_4\end{pmatrix} \tag{4i. 1. 1}$$

はg逆行列の1つであるとする．そうすると次の結果が成立する．

（a）　$\mathbf{XC}_2'\mathbf{X}=\mathbf{X},\ \mathbf{XC}_3\mathbf{X}=\mathbf{X}$　(4i. 1. 2)

（b）　$\mathbf{XC}_4\mathbf{X}'=\mathbf{XC}_4'\mathbf{X}'=\mathbf{GC}_3'\mathbf{X}'=\mathbf{XC}_3\mathbf{G}=\mathbf{GC}_2\mathbf{X}'=\mathbf{XC}_2'\mathbf{G}$　(4i. 1. 3)

（c）　$\mathbf{X}'\mathbf{C}_1\mathbf{X},\ \mathbf{X}'\mathbf{C}_1\mathbf{G},\ \mathbf{GC}_1\mathbf{X}$ はすべて零行列である．　(4i. 1. 4)

（d）　$\mathbf{GC}_1\mathbf{GC}_1\mathbf{G}=\mathbf{GC}_1\mathbf{G}=\mathbf{GC}_1'\mathbf{GC}_1\mathbf{G}=\mathbf{GC}_1'\mathbf{G}$　(4i. 1. 5)

（e）　$\mathrm{Tr}\,\mathbf{GC}_1=R(\mathbf{G}\vdots\mathbf{X})-R(\mathbf{X})=\mathrm{Tr}\,\mathbf{GC}_1'$　(4i. 1. 6)

（f）　$\mathbf{GC}_1\mathbf{G},\ \mathbf{XC}_4\mathbf{X}'$ は $\mathbf{C}_1,\mathbf{C}_4$ の選びかたにたいして不変である．　(4i. 1. 7)

方程式

$$\mathbf{Ga}+\mathbf{Xb}=\mathbf{0},\qquad \mathbf{X}'\mathbf{a}=\mathbf{X}'\mathbf{d} \tag{4i. 1. 8}$$

は，任意の $\mathbf{d}$ にたいして解をもち，その場合には，(4i. 1. 1) を用いると

$$\mathbf{a}=\mathbf{C}_2\mathbf{X}'\mathbf{d},\qquad \mathbf{b}=-\mathbf{C}_4\mathbf{X}'\mathbf{d} \tag{4i. 1. 9}$$

が1つの解であることがわかる．(4i. 1. 9) を (4i. 1. 8) に代入すると

$$\mathbf{GC}_2\mathbf{X}'\mathbf{d}-\mathbf{XC}_4\mathbf{X}'\mathbf{d}=\mathbf{0},\quad \mathbf{X}'\mathbf{C}_2\mathbf{X}'\mathbf{d}=\mathbf{X}'\mathbf{d},\quad \text{すべての } \mathbf{d} \text{ にたいして}$$

$$\Rightarrow \mathbf{GC}_2\mathbf{X}'=\mathbf{XC}_4\mathbf{X}'\quad \text{そして}\quad \mathbf{X}'\mathbf{C}_2\mathbf{X}'=\mathbf{X}'. \tag{4i. 1. 10}$$

$$\begin{pmatrix}\mathbf{G} & \mathbf{X}\\ \mathbf{X}' & \mathbf{0}\end{pmatrix}\begin{pmatrix}\mathbf{C}_1' & \mathbf{C}_3'\\ \mathbf{C}_2' & -\mathbf{C}_4'\end{pmatrix}\begin{pmatrix}\mathbf{G} & \mathbf{X}\\ \mathbf{X}' & \mathbf{0}\end{pmatrix}=\begin{pmatrix}\mathbf{G} & \mathbf{X}\\ \mathbf{X}' & \mathbf{0}\end{pmatrix}$$

であるから，(4i. 1. 1) の右辺の行列の転置はまた1つのg逆行列である．それゆえ，(4i. 1. 10) において $\mathbf{C}_2$ を $\mathbf{C}_3'$，$\mathbf{C}_4$ を $\mathbf{C}_4'$ でおきかえると

$$\mathbf{GC}_3'\mathbf{X}'=\mathbf{XC}_4'\mathbf{X}'\quad \text{そして}\quad \mathbf{X}'\mathbf{C}_3'\mathbf{X}'=\mathbf{X}' \tag{4i. 1. 11}$$

が得られる．(4i. 1. 11) の最初の方程式の両辺に $\mathbf{XC}_3$ をかけると

$$\mathbf{XC}_3\mathbf{GC}_3'\mathbf{X}'=\mathbf{XC}_3\mathbf{XC}_4'\mathbf{X}'=\mathbf{XC}_4'\mathbf{X}' \tag{4i. 1. 12}$$

となり，したがって $\mathbf{XC}_4'\mathbf{X}'$ は対称である．かくして (4i. 1. 10) から (4i. 1. 12) までは結果 (a)，(b) を証明している．

(c) を証明するためには，方程式

$$\mathbf{Ga}+\mathbf{Xb}=\mathbf{Xd},\qquad \mathbf{X}'\mathbf{a}=\mathbf{0} \tag{4i. 1. 13}$$

は任意の $\mathbf{d}$ にたいして解をもつことを確かめる．そうすると

$$\mathbf{a}=\mathbf{C}_1\mathbf{Xd},\qquad \mathbf{b}=\mathbf{C}_3\mathbf{Xd} \tag{4i. 1. 14}$$

が1つの解である．(4i. 1. 14) を (4i. 1. 13) に代入し，$\mathbf{d}$ を落すと

$$\mathbf{GC}_1\mathbf{X}+\mathbf{XC}_3\mathbf{X}=\mathbf{X},\qquad \mathbf{X}'\mathbf{C}_1\mathbf{X}=\mathbf{0}. \tag{4i. 1. 15}$$

しかし $\mathbf{XC}_3\mathbf{X}=\mathbf{X}$，かくして (c) が証明された．次に方程式

$$\mathbf{Ga}+\mathbf{Xb}=\mathbf{Gd},\qquad \mathbf{X}'\mathbf{a}=\mathbf{0} \tag{4i. 1. 16}$$

は任意の $\mathbf{d}$ にたいして解をもつことがわかる．1つの解は

$$\mathbf{a}=\mathbf{C}_1\mathbf{Gd},\qquad \mathbf{b}=\mathbf{C}_3\mathbf{Gd} \tag{4i. 1. 17}$$

である．(4i. 1. 17) を (4i. 1. 16) に代入し，$\mathbf{d}$ を落す．$\mathbf{GC}_1\mathbf{X}=\mathbf{0}$ であるから

$$\mathbf{GC}_1\mathbf{G}+\mathbf{XC}_3\mathbf{G}=\mathbf{G},\qquad \mathbf{X}'\mathbf{C}_1\mathbf{G}=\mathbf{0} \tag{4i. 1. 18}$$

$$\Rightarrow \mathbf{GC}_1\mathbf{GC}_1\mathbf{G}+\mathbf{GC}_1\mathbf{XC}_3\mathbf{G}=\mathbf{GC}_1\mathbf{G}=\mathbf{GC}_1\mathbf{GC}_1\mathbf{G}. \tag{4i. 1. 19}$$

また (4i. 1. 18) $\Rightarrow \mathbf{GC}_1'\mathbf{GC}_1\mathbf{G}=\mathbf{GC}_1\mathbf{G}$，これで (d) が証明された．さて，結果 $R(\mathbf{AA}^-)=R(\mathbf{A})=\mathrm{Tr}(\mathbf{AA}^-)$ を用いる [(iv)，**1b. 5**，(i)，**1b. 7**]．ただし $\mathbf{A}^-$ は $\mathbf{A}$ の任意のg逆行列である．

$$\begin{aligned}R\begin{pmatrix}\mathbf{G} & \mathbf{X}\\ \mathbf{X}' & \mathbf{0}\end{pmatrix}\begin{pmatrix}\mathbf{C}_1 & \mathbf{C}_2\\ \mathbf{C}_3 & -\mathbf{C}_4\end{pmatrix}&=\mathrm{Tr}\begin{pmatrix}\mathbf{GC}_1+\mathbf{XC}_3 & \mathbf{GC}_2-\mathbf{XC}_4\\ \mathbf{X}'\mathbf{C}_1 & \mathbf{X}'\mathbf{C}_2\end{pmatrix}\\ &=\mathrm{Tr}(\mathbf{GC}_1+\mathbf{XC}_3)+\mathrm{Tr}(\mathbf{X}'\mathbf{C}_2)\\ &=\mathrm{Tr}(\mathbf{GC}_1)+R(\mathbf{XC}_3)+R(\mathbf{X}'\mathbf{C}_2)\\ &=\mathrm{Tr}(\mathbf{GC}_1)+R(\mathbf{X})+R(\mathbf{X}')\end{aligned} \tag{4i. 1. 20}$$

が得られる．しかし

$$R\begin{pmatrix}\mathbf{G} & \mathbf{X}\\ \mathbf{X}' & \mathbf{0}\end{pmatrix}=R(\mathbf{G}\vdots\mathbf{X})+R(\mathbf{X}). \tag{4i. 1. 21}$$

(4i. 1. 20) と (4i. 1. 21) を等しいとおくと結果 (e) が得られる．

結果 (f) は，結果 (4i. 1. 2)—(4i. 1. 6) を用いると定義から得られる．

（ii）　$\begin{pmatrix}\mathbf{G} & \mathbf{X}\\ \mathbf{X}' & \mathbf{0}\end{pmatrix}$ が正則であるための必要十分条件は $R\begin{pmatrix}\mathbf{G}\\ \mathbf{X}'\end{pmatrix}=n$ かつ $R(\mathbf{X})=m$.

（iii）　$\mathbf{U}$は，$\mathscr{M}(\mathbf{X})\subset\mathscr{M}(\mathbf{G}+\mathbf{XUX}')$ かつ $\mathscr{M}(\mathbf{G})\subset\mathscr{M}(\mathbf{G}+\mathbf{XUX}')$ であるような行列とする．そうすると

$$\begin{pmatrix}\mathbf{G} & \mathbf{X}\\ \mathbf{X}' & \mathbf{0}\end{pmatrix}^{-}=\begin{pmatrix}\mathbf{C}_1 & \mathbf{C}_2\\ \mathbf{C}_3 & -\mathbf{C}_4\end{pmatrix}\Leftrightarrow\begin{pmatrix}\mathbf{G}+\mathbf{XUX}' & \mathbf{X}\\ \mathbf{X}' & \mathbf{0}\end{pmatrix}^{-}=\begin{pmatrix}\mathbf{C}_1 & \mathbf{C}_2\\ \mathbf{C}_3 & -\mathbf{U}-\mathbf{C}_4\end{pmatrix}.$$

命題 (ii)，(iii) の証明は容易である．

（iv）　(4i. 1. 1) における $\mathbf{C}_1,\mathbf{C}_2,\mathbf{C}_3,\mathbf{C}_4$ の具体的表現

（a）　$\mathscr{M}(\mathbf{X})\subset\mathscr{M}(\mathbf{G})$ ならば，$\mathbf{C}_1,\mathbf{C}_2,\mathbf{C}_3,\mathbf{C}_4$ の1つの選定は

$$\mathbf{C}_1=\mathbf{G}^--\mathbf{G}^-\mathbf{XC}_3,\qquad \mathbf{C}_3=(\mathbf{X}'\mathbf{G}^-\mathbf{X})^-\mathbf{X}'\mathbf{G}^-$$

$$\mathbf{C}_2=\mathbf{C}_3',\qquad \mathbf{C}_4=(\mathbf{X}'\mathbf{G}^-\mathbf{X})^-$$

である．

（b）$R(\mathbf{G})=n$ ならば，$\mathbf{C}_1,\mathbf{C}_2,\mathbf{C}_3,\mathbf{C}_4$ の1つの選定は

$$\mathbf{C}_1=\mathbf{G}^{-1}-\mathbf{G}^{-1}\mathbf{X}\mathbf{C}_3,\qquad \mathbf{C}_3=(\mathbf{X}'\mathbf{G}^{-1}\mathbf{X})^{-}\mathbf{X}'\mathbf{G}^{-1}$$
$$\mathbf{C}_2=\mathbf{C}_3',\qquad \mathbf{C}_4=(\mathbf{X}'\mathbf{G}^{-1}\mathbf{X})^{-}$$

である．

（c）一般に，$\mathbf{C}_1,\mathbf{C}_2,\mathbf{C}_3,\mathbf{C}_4$ の1つの選定は

$$\mathbf{C}_1=\mathbf{W}-\mathbf{W}\mathbf{X}\mathbf{C}_3,\qquad \mathbf{C}_3=\mathbf{T}\mathbf{X}'\mathbf{W},$$
$$\mathbf{C}_2=\mathbf{C}_3',\qquad \mathbf{C}_4=-\mathbf{U}+\mathbf{T}$$

である．ただし $\mathbf{T}=(\mathbf{X}'(\mathbf{G}+\mathbf{X}'\mathbf{U}\mathbf{X})^{-}\mathbf{X})^{-}$，$\mathbf{W}=(\mathbf{G}+\mathbf{X}'\mathbf{U}\mathbf{X})^{-}$ であり，$\mathbf{U}$ は $\mathscr{M}(\mathbf{X})\subset\mathscr{M}(\mathbf{G}+\mathbf{X}\mathbf{U}\mathbf{X}')$ かつ $\mathscr{M}(\mathbf{G})\subset\mathscr{M}(\mathbf{G}+\mathbf{X}\mathbf{U}\mathbf{X}')$ であるような任意の行列である．

命題（iv）は直接検証するか，または次の型の方程式

$$\mathbf{G}\mathbf{a}+\mathbf{X}\mathbf{b}=\boldsymbol{\alpha}$$
$$\mathbf{X}'\mathbf{a}=\boldsymbol{\beta}$$

を普通の消去法で解くことにより証明できる．

4i.2　一般ガウス＝マルコフ模型（GGM）

さて，一般ガウス＝マルコフ（GGM）模型

$$(\mathbf{Y},\mathbf{X}\boldsymbol{\beta},\sigma^2\mathbf{G})\tag{4i.2.1}$$

を考えよう．ただし $\mathrm{R}(\mathbf{X})$ に不足があってもよいし，$\mathbf{G}$ は特異であってもよい．GGM 模型は母数 $\boldsymbol{\beta}$ に指定された制約がある場合を含むということを注意しておく．かくして，もし $\mathbf{R}\boldsymbol{\beta}=\mathbf{C}$ が制約式の組であるならば

$$\mathbf{Y}_e=\begin{pmatrix}\mathbf{X}\\ \mathbf{C}\end{pmatrix},\qquad \mathbf{X}_e=\begin{pmatrix}\mathbf{X}\\ \mathbf{R}\end{pmatrix}\tag{4i.2.2}$$

を考えることができ，この場合には

$$\sigma^2\mathbf{G}_e=D(\mathbf{Y}_e)=\sigma^2\begin{pmatrix}\mathbf{G}&0\\0&0\end{pmatrix}\tag{4i.2.3}$$

であって，拡張した模型を

$$(\mathbf{Y}_e,\mathbf{X}_e\boldsymbol{\beta},\sigma^2\mathbf{G}_e)\tag{4i.2.4}$$

と，(4i.2.1) と同じ形に書くことができる．

$\mathbf{G}$ が特異であるときには，従属変数 $\mathbf{Y}$ についてある自然な制約があり，模型に明らかな不一致がないということを確かめなければいけない．

（i）**一致性の検定**　$\mathbf{Y}\in\mathscr{M}(\mathbf{G}\,\vdots\,\mathbf{X})$.

$\boldsymbol{\alpha}$ は $\boldsymbol{\alpha}'\mathbf{X}=0$ かつ $\boldsymbol{\alpha}'\mathbf{G}=0$ であるようなベクトルとする．そうすると明らかに $E(\boldsymbol{\alpha}'\mathbf{Y})=0$ かつ $E(\boldsymbol{\alpha}'\mathbf{Y})^2=0\Rightarrow\boldsymbol{\alpha}'\mathbf{Y}=0\Rightarrow$ 確率1で $\mathbf{Y}\in\mathscr{M}(\mathbf{G}\,\vdots\,\mathbf{X})$.

（ii）n 次元確率ベクトル $\mathbf{Y}$ と未知ベクトル $\boldsymbol{\beta}$ は $(\mathbf{Y}-\mathbf{X}\boldsymbol{\beta})\in\mathscr{M}(\mathbf{G})$ を満足するものである．

$\boldsymbol{\alpha}'\mathbf{G}=0\Rightarrow\boldsymbol{\alpha}'(\mathbf{Y}-\mathbf{X}\boldsymbol{\beta})=0$ よりこの結果が得られる．かくして $\mathbf{G}$ の特異性は $\mathbf{Y}$ と $\boldsymbol{\beta}$ にある制約を課す．

$\mathbf{G}$ が特異ならば，確率1で $\mathbf{H}\mathbf{Y}=0$ であるような行列 $\mathbf{H}$ が存在し，これは，$\mathbf{H}\mathbf{X}\boldsymbol{\beta}=0$ が $\boldsymbol{\beta}$ についての制約であるということを導く．もちろん，$\mathbf{Y}$ が既知のときに限り $\mathbf{H}$ が既知である．$\mathbf{H}$ が既知のときには，$\mathbf{L}'\mathbf{Y}$ が $\mathbf{p}'\boldsymbol{\beta}$ にたいして不偏であるための条件 $\mathbf{X}'\mathbf{L}=\mathbf{p}$ は不必要である．

（iii）$\mathbf{L}'\mathbf{Y}$ が $\mathbf{p}'\boldsymbol{\beta}$ にたいして不偏であるための必要十分条件は，$\mathbf{X}'(\mathbf{L}+\mathbf{H}'\boldsymbol{\lambda})=\mathbf{p}$ であるようなベクトル $\boldsymbol{\lambda}$ が存在することである．ただし，$\mathbf{H}$ は $\mathbf{H}\mathbf{Y}=0$ であるような最大階数の行列である．しかしながら，$\mathbf{L}'\mathbf{Y}$ が $\mathbf{p}'\boldsymbol{\beta}$ にたいして不偏であるような $\mathbf{L}$ が与えられたとき，$\mathbf{M}'\mathbf{Y}=\mathbf{L}'\mathbf{Y}$ かつ $\mathbf{X}'\mathbf{M}=\mathbf{p}$ であるような $\mathbf{M}$ を見いだすことができる．$R(\mathbf{H})=n-R(\mathbf{G})$ または $n-R(\mathbf{G})-1$ であることを注意しておく．

容易に証明されるこの結果は重要である．それは，$\mathbf{p}'\boldsymbol{\beta}$ の BLUE を捜すのに，$\mathbf{X}'\mathbf{L}=\mathbf{p}$ であるような線形関数 $\mathbf{L}'\mathbf{Y}$ だけを考えるか，またはあたかも $\boldsymbol{\beta}$ が自由な母数のように考えていけばよいということを示している．またもし $\mathbf{L}'\mathbf{Y}$ の期待値が零であるならば，$\mathbf{L}'\mathbf{Y}=\mathbf{K}'\mathbf{Y}$ かつ $\mathbf{K}'\mathbf{X}=0$ であるようなベクトル $\mathbf{K}$ を見いだすことができる．

4i.3　逆分割行列（IPM）法

GGM 模型 (4i.2.1) を考え，未知母数の推測問題が，どのようにして g 逆行列

$$\begin{pmatrix}\mathbf{G}&\mathbf{X}\\ \mathbf{X}'&0\end{pmatrix}^{-}=\begin{pmatrix}\mathbf{C}_1&\mathbf{C}_2\\ \mathbf{C}_3&-\mathbf{C}_4\end{pmatrix}\tag{4i.3.1}$$

の数値評価に帰着されるかを示そう（Rao (1971e, 1972c)）．もっと正確には，適当なアルゴリズム（計算プログラム）によって (4i.3.1) の左辺の行列が逆転されて $\mathbf{C}_1,\mathbf{C}_2,\mathbf{C}_3,\mathbf{C}_4$ が求められるならば，BLUE，それらの分散・共分散，σ^2 の不偏推定量，そして線形仮説を検定するための検定基準を求めるために必要なものすべてを得ている．命題 (i), (ii) は $\mathbf{G}$ が特異であってもなくても成立する基礎的な結果を含む．

（i）$\mathbf{C}_1,\mathbf{C}_2,\mathbf{C}_3,\mathbf{C}_4$ は (4i.3.1) で定義されたものであるとする．そうすると次のことが成立する．

（a）（$\mathbf{C}_2$ または $\mathbf{C}_3$ の利用）　推定可能な（すなわち，不偏推定量をもつ）母数の関数 $\mathbf{p}'\boldsymbol{\beta}$ の BLUE は $\mathbf{p}'\hat{\boldsymbol{\beta}}$ である．ただし

$$\hat{\boldsymbol{\beta}}=\mathbf{C}_3\mathbf{Y}\quad\text{または}\quad\hat{\boldsymbol{\beta}}=\mathbf{C}_2'\mathbf{Y}.\tag{4i.3.2}$$

（b）（$\mathbf{C}_4$ の利用）　$\hat{\beta}$ の分散共分散行列は，$\mathbf{p}'\beta, \mathbf{q}'\beta$ が推定可能であるときにはいつでも

$$V(\mathbf{p}'\hat{\beta}) = \sigma^2\mathbf{p}'\mathbf{C}_4\mathbf{p}$$
$$\mathrm{cov}(\mathbf{p}'\hat{\beta}, \mathbf{q}'\hat{\beta}) = \sigma^2\mathbf{p}'\mathbf{C}_4\mathbf{q} = \sigma^2\mathbf{q}'\mathbf{C}_4\mathbf{p} \qquad (4\mathrm{i}.3.3)$$

という意味において，$\sigma^2\mathbf{C}_4$ である．

（c）（$\mathbf{C}_1$ の利用）　σ^2 の不偏推定量は

$$\hat{\sigma}^2 = f^{-1}\mathbf{Y}'\mathbf{C}_1\mathbf{Y} \qquad (4\mathrm{i}.3.4)$$

である．ただし $f = R(\mathbf{G}\vdots\mathbf{X}) - R(\mathbf{X})$.

(a) の証明　$\mathbf{L}'\mathbf{Y}$ が $E(\mathbf{L}'\mathbf{Y}) = \mathbf{p}'\beta$ であるならば $\mathbf{X}'\mathbf{L} = \mathbf{p}$ である．この条件のもとで $V(\mathbf{L}'\mathbf{Y}) = \sigma^2\mathbf{L}'\mathbf{G}\mathbf{L}$ または $\mathbf{L}'\mathbf{G}\mathbf{L}$ を最小にしなければいけない．$\mathbf{L}_*$ を最適な選定とし，$\mathbf{L}$ を $\mathbf{X}'\mathbf{L} = \mathbf{X}'\mathbf{L}_*$ であるような任意の他のベクトルとする．そうすると

$$\begin{aligned}\mathbf{L}'\mathbf{G}\mathbf{L} &= (\mathbf{L} - \mathbf{L}_* + \mathbf{L}_*)'\mathbf{G}(\mathbf{L} - \mathbf{L}_* + \mathbf{L}_*)\\ &= (\mathbf{L} - \mathbf{L}_*)'\mathbf{G}(\mathbf{L} - \mathbf{L}_*) + \mathbf{L}_*'\mathbf{G}\mathbf{L}_* + 2\mathbf{L}_*'\mathbf{G}(\mathbf{L} - \mathbf{L}_*) \geqslant \mathbf{L}_*'\mathbf{G}\mathbf{L}_*\end{aligned}$$

$\Leftrightarrow \mathbf{X}'(\mathbf{L} - \mathbf{L}_*) = \mathbf{0}$ のときはいつでも $\mathbf{L}_*'\mathbf{G}(\mathbf{L} - \mathbf{L}_*) = 0$，すなわち適当な $\mathbf{K}_*$ にたいして $\mathbf{G}\mathbf{L}_* = -\mathbf{X}\mathbf{K}_*$

したがって方程式

$$\mathbf{G}\mathbf{L}_* + \mathbf{X}\mathbf{K}_* = 0,$$
$$\mathbf{X}'\mathbf{L}_* = \mathbf{p} \qquad (4\mathrm{i}.3.5)$$

をみたす $\mathbf{L}_*, \mathbf{K}_*$ が存在する．(4i. 3. 5) は解をもつから

$$\mathbf{L}_* = \mathbf{C}_2\mathbf{p} \quad \text{または} \quad \mathbf{C}_3'\mathbf{p}, \qquad \mathbf{K}_* = -\mathbf{C}_4\mathbf{p}$$

が1つの解である．したがって $\mathbf{p}'\beta$ の BLUE は $\mathbf{L}_*'\mathbf{Y} = \mathbf{p}'\mathbf{C}_2'\mathbf{Y} = \mathbf{p}'\mathbf{C}_3\mathbf{Y}$ である．

(b) の証明　ある $\mathbf{M}$ にたいして $\mathbf{p} = \mathbf{X}'\mathbf{M}$ という事実を使う．そうすると

$$\begin{aligned}V(\mathbf{p}'\mathbf{C}_2'\mathbf{Y}) &= \sigma^2\mathbf{M}'(\mathbf{X}\mathbf{C}_2'\mathbf{G})\mathbf{C}_2\mathbf{X}'\mathbf{M}\\ &= \sigma^2\mathbf{M}'\mathbf{X}\mathbf{C}_4(\mathbf{X}'\mathbf{C}_2\mathbf{X}')\mathbf{M} \qquad (4\mathrm{i}.1.3)\ \text{を用いた}\\ &= \sigma^2\mathbf{M}'\mathbf{X}\mathbf{C}_4\mathbf{X}'\mathbf{M} \qquad (4\mathrm{i}.1.2)\ \text{を用いた}\\ &= \sigma^2\mathbf{p}'\mathbf{C}_4\mathbf{p}.\end{aligned}$$

同様にして

$$\mathrm{cov}(\mathbf{p}'\mathbf{C}_2'\mathbf{Y}, \mathbf{q}'\mathbf{C}_2'\mathbf{Y}) = \sigma^2\mathbf{p}'\mathbf{C}_4\mathbf{q} = \sigma^2\mathbf{q}'\mathbf{C}_4\mathbf{p}.$$

(c) の証明　(4i. 1. 4) を用いると，$\mathbf{X}'\mathbf{C}_1\mathbf{G} = 0$，$\mathbf{X}'\mathbf{C}_1\mathbf{X} = 0$ であるから

$$\mathbf{Y}'\mathbf{C}_1\mathbf{Y} = (\mathbf{Y} - \mathbf{X}\beta)'\mathbf{C}_1(\mathbf{Y} - \mathbf{X}\beta)$$
$$\begin{aligned}E[(\mathbf{Y} - \mathbf{X}\beta)'\mathbf{C}_1(\mathbf{Y} - \mathbf{X}\beta)] &= \mathrm{Tr}\,\mathbf{C}_1E[(\mathbf{Y} - \mathbf{X}\beta)(\mathbf{Y} - \mathbf{X}\beta)']\\ &= \sigma^2\mathrm{Tr}\,\mathbf{C}_1\mathbf{G} = \sigma^2[R(\mathbf{G}\vdots\mathbf{X}) - R(\mathbf{X})], \qquad (4\mathrm{i}.1.6)\ \text{を用いた}\end{aligned}$$

（ii）$\mathbf{P}'\hat{\beta}$ は k 個の推定可能な母数の関数の組 $\mathbf{P}'\beta$ の BLUE のベクトルとし，$R_0^2 = \mathbf{Y}'\mathbf{C}_1\mathbf{Y}$，$f$ は（i）で定義されたものとする．

$\mathbf{Y} \sim N_n(\mathbf{X}\beta, \sigma^2\mathbf{G})$ ならば次の結果が得られる：

（a）$\mathbf{P}'\hat{\beta}$ と $\mathbf{Y}'\mathbf{C}_1\mathbf{Y}$ は独立に分布し，

$$\mathbf{P}'\hat{\beta} \sim N_k(\mathbf{P}'\beta, \sigma^2\mathbf{D}), \qquad \mathbf{D} = \mathbf{P}'\mathbf{C}_4\mathbf{P} \qquad (4\mathrm{i}.3.6)$$
$$R_0^2 = \mathbf{Y}'\mathbf{C}_1\mathbf{Y} \sim \sigma^2\chi_f^2. \qquad (4\mathrm{i}..7)$$

（b）$\mathbf{P}'\beta = \mathbf{w}$ を帰無仮説とする．この帰無仮説が意味をもつ必要十分条件は

$$\mathbf{D}\mathbf{D}^-\mathbf{u} = \mathbf{u}, \quad \text{ただし} \quad \mathbf{u} = \mathbf{P}'\hat{\beta} - \mathbf{w} \qquad (4\mathrm{i}.3.8)$$

である．もしこの仮説が意味をもつならば，

$$F = \frac{\mathbf{u}'\mathbf{D}^-\mathbf{u}}{h} \div \frac{R_0^2}{f}, \qquad h = R(\mathbf{D}) \qquad (4\mathrm{i}.3.9)$$

は，仮説が正しいときには自由度 h, f の中心 F 分布に従い，そうでないときには非心分布に従う．

(a) の証明　$E(\mathbf{P}'\hat{\beta}) = \mathbf{P}'\beta$，$D(\mathbf{P}'\hat{\beta}) = \sigma^2\mathbf{P}'\mathbf{C}_4\mathbf{P}$ であるから，(3b. 3. 5) を用いると，$\mathbf{P}'\hat{\beta} \sim N_k(\mathbf{P}'\beta, \sigma^2\mathbf{P}'\mathbf{C}_4\mathbf{P})$.

$\mathbf{F} = (\mathbf{C}_1 + \mathbf{C}_1')/2$ とする．そうすると，(4i. 1. 5) を用いると $\mathbf{Y}'\mathbf{C}_1\mathbf{Y} = \mathbf{Y}'\mathbf{F}\mathbf{Y}$ そして $\mathbf{G}\mathbf{F}\mathbf{G}\mathbf{F}\mathbf{G} = \mathbf{G}\mathbf{F}\mathbf{G}$. したがって (3b. 4. 7) より $\mathbf{Y}'\mathbf{C}_1\mathbf{Y}$ は $\sigma^2\chi^2$ 分布をもつ．

$\mathscr{M}(\mathbf{P}) \subset \mathscr{M}(\mathbf{X}')$ であるから $\mathbf{P}'\hat{\beta} = \mathbf{P}'\mathbf{C}_2'\mathbf{Y} = \mathbf{H}'\mathbf{X}\mathbf{C}_2'\mathbf{Y}$. したがって (4i. 1. 3) を用いると $\mathbf{G}\mathbf{F}\mathbf{G}\mathbf{C}_2\mathbf{X}'\mathbf{H} = 0$，これは (3b. 4. 9) より，$\mathbf{Y}'\mathbf{C}_1\mathbf{Y}$ と $\mathbf{P}'\hat{\beta}$ が独立に分布することを示す．[$\mathbf{G}\mathbf{F}\mathbf{X} = 0$ であるから $\mathbf{G}\mathbf{F}(\mathbf{G}\mathbf{C}_2\mathbf{X}')\mathbf{H} = \mathbf{G}\mathbf{F}(\mathbf{X}\mathbf{C}_2'\mathbf{G}\mathbf{H}) = 0$ ということに注意せよ．]

(b) の証明　$V(\alpha'\mathbf{P}'\hat{\beta}) = 0 \Rightarrow \alpha'(\mathbf{P}'\hat{\beta} - \mathbf{w}) = 0 \Rightarrow \mathbf{P}'\hat{\beta} - \mathbf{w} \in \mathscr{M}(\mathbf{D})$，ならば仮説は意味をもつ．これの必要十分条件は $\mathbf{u} = \mathbf{P}'\hat{\beta} - \mathbf{w}$ とするとき $\mathbf{D}\mathbf{D}^-\mathbf{u} = \mathbf{u}$ である．もし仮説が正しいならば，$\mathbf{D}\mathbf{D}^-\mathbf{D} = \mathbf{D}$ であるから，$\mathbf{u}'\mathbf{D}^-\mathbf{u}$ は自由度 $h = R(\mathbf{D})$ の中心 $\sigma^2\chi^2$ 分布をもつ．$\mathbf{u}$ と R_0^2 は独立に分布するから，$\mathbf{u}'\mathbf{D}^-\mathbf{u}$ と R_0^2 は独立であり，したがって統計量 (4i. 3. 9) は自由度 h, f の F 分布をもつ．もし仮説が正しくないならば，この分布は非心であろう．

4i. 4　最小2乗法の統一理論

模型 $(\mathbf{Y}, \mathbf{X}\beta, \sigma^2\mathbf{G})$ を考えよう．$|\mathbf{G}| \neq 0$ のときには，β を推定するための正規方程式は

$$(\mathbf{Y} - \mathbf{X}\beta)'\mathbf{G}^{-1}(\mathbf{Y} - \mathbf{X}\beta)$$

の β に関する微分を零ベクトルに等しいとおくことにより得られるということを 4a. 7（表 4a. 7 参照）で示した．この手順は $|\mathbf{G}| = 0$ のときには使えない．すべての場合に適用可能で最小2乗法と異なる1つの一般的方法を 4i. 3 で与えた．すべての場合に適用可能な最

小2乗法の類似が展開可能かどうかを調べよう．もっと明確には，$|\mathbf{G}|=0$ であっても $\neq 0$ であっても，β の推定値 $\hat{\beta}$ は

$$(\mathbf{Y}-\mathbf{X}\beta)'\mathbf{M}(\mathbf{Y}-\mathbf{X}\beta)$$

の停滞値，すなわち β に関する微分が零になる値であり，σ^2 の推定値は

$$(\mathbf{Y}-\mathbf{X}\hat{\beta})'\mathbf{M}(\mathbf{Y}-\mathbf{X}\hat{\beta})\div f$$

として求められるような行列 $\mathbf{M}$ を見いだすことができるか？　ただし (4i. 3. 4) のように $f=R(\mathbf{G}\vdots\mathbf{X})-R(\mathbf{X})$.

$\mathbf{M}$ の1つの選定は任意の g 逆行列 $(\mathbf{G}+\mathbf{XUX}')^-$ であることを示す．ただし $\mathbf{U}$ は，$R(\mathbf{G}\vdots\mathbf{X})=R(\mathbf{G}+\mathbf{XUX}')$ であるような任意の対称行列であり，これの必要性は Rao (1973a) に示されている．Rao and Mitra (1971h) に示されているように，$\mathbf{U}$ の1つの簡単な選定は $\mathbf{U}=k^2\mathbf{I}$, $k\neq 0$ である．$\mathscr{M}(\mathbf{X})\subset\mathscr{M}(\mathbf{G})$ のときには k を零にとることができる．$\mathbf{T}=\mathbf{G}+\mathbf{XUX}'$ と書き，それの任意の g 逆行列（必ずしも対称ではない）を $\mathbf{T}^-$ と書く．

（i）$(\mathbf{Y},\mathbf{X}\beta,\sigma^2\mathbf{G})$ を GGM 模型（すなわち，$\mathbf{G}$ は特異であってもよいし，$\mathbf{X}$ に階数おちがあってもよい）とし，$\mathbf{T},\mathbf{T}^-,\mathbf{U}$ は上に定義したものとする．

（a）推定可能な関数 $\mathbf{p}'\beta$ の BLUE は $\mathbf{p}'\hat{\beta}$ である．ここで $\hat{\beta}$ は $(\mathbf{Y}-\mathbf{X}\beta)'\mathbf{T}^-(\mathbf{Y}-\mathbf{X}\beta)$ の停滞値である．

（b）$V(\mathbf{p}'\hat{\beta})=\sigma^2\mathbf{p}'[(\mathbf{X}'\mathbf{T}^-\mathbf{X})^--\mathbf{U}]\mathbf{p}$,

$$\mathrm{cov}(\mathbf{p}'\hat{\beta},\mathbf{q}'\hat{\beta})=\sigma^2\mathbf{p}'[(\mathbf{X}'\mathbf{T}^-\mathbf{X})^--\mathbf{U}]\mathbf{q} \tag{4i. 4. 1}$$

（c）σ^2 の1つの不偏推定量は

$$\hat{\sigma}^2=f^{-1}(\mathbf{Y}-\mathbf{X}\hat{\beta})'\mathbf{T}^-(\mathbf{Y}-\mathbf{X}\hat{\beta})=f^{-1}R_0^2 \tag{4i. 4. 2}$$

である．ただし $f=R(\mathbf{G}\vdots\mathbf{X})-R(\mathbf{X})$.

(a) の証明　$\mathbf{X},\mathbf{Y}\in\mathscr{M}(\mathbf{T})$ で $\mathbf{T}$ は対称であるから，$\mathbf{X}'\mathbf{T}^-\mathbf{X}=\mathbf{X}'(\mathbf{T}^-)'\mathbf{X}$, $\mathbf{X}'\mathbf{T}^-\mathbf{Y}=\mathbf{X}'(\mathbf{T}^-)'\mathbf{Y}$ に注意すると，$\hat{\beta}$ は方程式

$$\mathbf{X}'\mathbf{T}^-\mathbf{X}\hat{\beta}=\mathbf{X}'\mathbf{T}^-\mathbf{Y} \tag{4i. 4. 3}$$

を満足する．$\hat{\beta}=(\mathbf{X}'\mathbf{T}^-\mathbf{X})^-\mathbf{X}'\mathbf{T}^-\mathbf{Y}$ を1つの解とする．71ページの問題30（第1章の終り）で示したように，$R(\mathbf{X}'\mathbf{T}^-\mathbf{X})=R(\mathbf{X}')$ であるから解は存在する．

$\mathbf{p}'\beta$ を推定可能な関数とする．すなわち，ある $\mathbf{L}$ にたいして $\mathbf{p}=\mathbf{X}'\mathbf{L}$. そうすると，[vi(b), 1b. 5] を使うと $\mathbf{X}'\in\mathscr{M}(\mathbf{X}'\mathbf{T}^-\mathbf{X})$ であるから

$$\mathbf{p}'\hat{\beta}=\mathbf{L}'\mathbf{X}(\mathbf{X}'\mathbf{T}^-\mathbf{X})^-\mathbf{X}'\mathbf{T}^-\mathbf{Y}$$

$$E(\mathbf{p}'\hat{\beta})=\mathbf{L}'\mathbf{X}(\mathbf{X}'\mathbf{T}^-\mathbf{X})^-\mathbf{X}'\mathbf{T}^-\mathbf{X}\beta=\mathbf{L}'\mathbf{X}\beta=\mathbf{p}'\beta.$$

さらに $\mathbf{Z}$ は $\mathbf{Z}'\mathbf{X}=0$ であるものとするならば，$\mathbf{X}\in\mathscr{M}(\mathbf{T})$ のため $\mathbf{X}'\mathbf{T}^-\mathbf{GZ}=\mathbf{X}'\mathbf{T}^-(\mathbf{G}+\mathbf{XUX}')\mathbf{Z}=\mathbf{X}'\mathbf{T}^-\mathbf{TZ}=\mathbf{X}'\mathbf{Z}=0$ であるから，

$$\mathrm{cov}(\mathbf{p}'\hat{\beta},\mathbf{Z}'\mathbf{Y})=\sigma^2\mathbf{L}'\mathbf{X}(\mathbf{X}'\mathbf{T}^-\mathbf{X})^-\mathbf{X}'\mathbf{T}^-\mathbf{GZ}=0.$$

よって $\mathbf{p}'\hat{\beta}$ は BLUE である．

(b) の証明

$$\begin{aligned}V(\mathbf{p}'\hat{\beta})&=V(\mathbf{L}'\mathbf{X}(\mathbf{X}'\mathbf{T}^-\mathbf{X})^-\mathbf{X}'\mathbf{T}^-\mathbf{Y})\\&=V(\mathbf{L}'\mathbf{WY}),\qquad \mathbf{W}=\mathbf{X}(\mathbf{X}'\mathbf{T}^-\mathbf{X})^-\mathbf{X}'\mathbf{T}^-\\&=\sigma^2\mathbf{L}'\mathbf{W}(\mathbf{G}+\mathbf{XUX}')\mathbf{W}'\mathbf{L}-\mathbf{L}'\mathbf{WXUX}'\mathbf{W}'\mathbf{L}\\&=\sigma^2[\mathbf{L}'\mathbf{X}(\mathbf{X}'\mathbf{T}^-\mathbf{X})^-\mathbf{X}'\mathbf{L}-\mathbf{L}'\mathbf{XUX}'\mathbf{L}]\\&=\sigma^2\mathbf{p}'[(\mathbf{X}'\mathbf{T}^-\mathbf{X})^--\mathbf{U}]\mathbf{p}.\end{aligned}$$

同様にして，共分散の表現も証明される．

(c) の証明　(4a. 5. 4) の線に沿って次のことが容易に示される．

$$\begin{aligned}E(\mathbf{Y}-\mathbf{X}\hat{\beta})(\mathbf{Y}-\mathbf{X}\hat{\beta})'&=D(\mathbf{Y})-D(\mathbf{X}\hat{\beta})\\&=\sigma^2\{\mathbf{G}-\mathbf{X}[(\mathbf{X}'\mathbf{T}^-\mathbf{X})^--\mathbf{U}]\mathbf{X}'\}\\&=\sigma^2\{\mathbf{T}-\mathbf{X}(\mathbf{X}'\mathbf{T}^-\mathbf{X})^-\mathbf{X}'\}\end{aligned}$$

$$\begin{aligned}\sigma^{-2}E(\mathbf{Y}-\mathbf{X}\hat{\beta})'\mathbf{T}^-(\mathbf{Y}-\mathbf{X}\hat{\beta})&=\mathrm{Tr}\,\mathbf{T}^-[\mathbf{T}-\mathbf{X}(\mathbf{X}'\mathbf{T}^-\mathbf{X})^-\mathbf{X}']\\&=\mathrm{Tr}\,\mathbf{T}^-\mathbf{T}-\mathrm{Tr}(\mathbf{X}'\mathbf{T}^-\mathbf{X})(\mathbf{X}'\mathbf{T}^-\mathbf{X})^-\\&=R(\mathbf{T})-R(\mathbf{X})=f\end{aligned}$$

よって (c) が証明された．

線形仮説の検定は，(4i. 3. 6) の $\mathbf{C}_4$ として $(\mathbf{X}'\mathbf{T}^-\mathbf{X})^--\mathbf{U}$, $\mathbf{Y}'\mathbf{C}_1\mathbf{Y}$ として (4i. 4. 2) で定義した R_0^2 をそれぞれとり，4i. 3 の命題 (ii) と同じ方法でやればよい．

$\mathbf{U}$ として，もし必要ならば，条件 $R(\mathbf{G}\vdots\mathbf{X})=R(\mathbf{T})$ に加えて $\mathbf{T}=\mathbf{G}+\mathbf{XUX}'$ が n. n. d. であるように選ぶことができるということを注意する．この場合には $\mathbf{T}^-$ の任意の選定にたいして $(\mathbf{Y}-\mathbf{X}\beta)'\mathbf{T}^-(\mathbf{Y}-\mathbf{X}\beta)$ は非負であり，この式は停滞点 $\hat{\beta}$ で最小に到達する．

一般には $\mathbf{T}^-$ は $\mathbf{G}$ の g 逆行列ではない．しかしながら，$\mathbf{G}$ と $\mathbf{XUX}'$ が実質的に素な空間をつくるように $\mathbf{U}$ を選ぶことができ，その場合には $\mathbf{T}^-$ は $\mathbf{G}$ の g 逆行列である．不幸にして任意の検定可能な仮説 $\mathbf{P}'\beta=\mathbf{w}$ にたいして，$\mathbf{u}'\mathbf{D}^-\mathbf{u}=\min_{\mathbf{P}'\beta=\mathbf{w}}(\mathbf{Y}-\mathbf{X}\beta)'\mathbf{M}(\mathbf{Y}-\mathbf{X}\beta)-\min_{\beta}(\mathbf{Y}-\mathbf{X}\beta)'\mathbf{M}(\mathbf{Y}-\mathbf{X}\beta)$ は $\mathbf{G}$ が特異であるときには真ではない．ただし $\mathbf{u}'\mathbf{D}^-\mathbf{u}$ は (4i. 3. 9) で定義されたものである．しかしながら $\mathscr{M}(\mathbf{X})\subset\mathscr{M}(\mathbf{G})$ でなければ $\mathbf{M}$ は選ばれる (Rao (1972d) を参照)．

4j　分散成分の推定

4j. 1　分散成分模型

4f において一般2元分類模型を導入し，分散成分の推定の特別な問題を考えた．さて一

般線形模型 $\mathbf{Y}=\mathbf{X}\boldsymbol{\beta}+\boldsymbol{\varepsilon}$ を考える．ただし誤差成分 $\boldsymbol{\varepsilon}$ は構造

$$\boldsymbol{\varepsilon}=\mathbf{U}_1\boldsymbol{\xi}_1+\cdots+\mathbf{U}_k\boldsymbol{\xi}_k \tag{4j.1.1}$$

をもつ．ここで $\boldsymbol{\xi}_i$ は

$$E(\boldsymbol{\xi}_i)=0,\qquad D(\boldsymbol{\xi}_i)=\sigma_i^2\mathbf{I},\qquad i=1,\cdots,k$$
$$E(\boldsymbol{\xi}_i\boldsymbol{\xi}_j')=0,\qquad i\neq j \tag{4j.1.2}$$

であるような仮説的な（観測不可能な）変量の n_i 次のベクトルである．そうすると，

$$D(\boldsymbol{\varepsilon})=\sigma_1^2\mathbf{U}_1\mathbf{U}_1'+\cdots+\sigma_k^2\mathbf{U}_k\mathbf{U}_k'=\sigma_1^2\mathbf{V}_1+\cdots+\sigma_k^2\mathbf{V}_k. \tag{4j.1.3}$$

計画行列 $\mathbf{X},\mathbf{U}_i$（または $\mathbf{V}_i$）と観測値ベクトル $\mathbf{Y}$ は既知であるが，$\boldsymbol{\beta}$ と $\boldsymbol{\sigma}=(\sigma_1^2,\cdots,\sigma_k^2)$は推定すべき未知母数である．仮説的な変量 $\boldsymbol{\xi}_1,\cdots,\boldsymbol{\xi}_k$ の分散 $\sigma_1^2,\cdots,\sigma_k^2$ は分散成分と呼ばれる．

$\alpha_1,\cdots,\alpha_k$ を $\sigma_1,\cdots,\sigma_k$ の事前の（または近似的な）値とする．（もし事前の情報がないならば，各 α_i は1としてよい．）そうすると (4j.1.1) と (4j.1.3) を

$$\boldsymbol{\varepsilon}=\mathbf{W}_1\boldsymbol{\eta}_1+\cdots+\mathbf{W}_k\boldsymbol{\eta}_k,\qquad \mathbf{W}_i=\alpha_i\mathbf{U}_i,\qquad \boldsymbol{\xi}_i=\alpha_i\boldsymbol{\eta}_i,$$
$$D(\boldsymbol{\varepsilon})=\gamma_1^2\mathbf{T}_1+\cdots+\gamma_k^2\mathbf{T}_k,\qquad \mathbf{T}_i=\mathbf{W}_i\mathbf{W}_i'=\alpha_i^2\mathbf{V}_i \tag{4j.1.4}$$

と書きかえることができる．ここで $\gamma_i^2=\sigma_i^2/\alpha_i^2$ は推定すべき尺度化された分散成分である．

4j.2　MINQUE 理論

分散成分の線形関数

$$\sum p_i\sigma_i^2=\sum q_i\gamma_i^2 \tag{4j.2.1}$$

を，次の条件のもとで2次関数 $\mathbf{Y}'\mathbf{A}\mathbf{Y}$ で推定することを考えよう．

（a）不変性と不偏性　$\boldsymbol{\beta}$ の代りに $\boldsymbol{\beta}_d=\boldsymbol{\beta}-\boldsymbol{\beta}_0$ を考えるならば，模型は $\mathbf{Y}_d=\mathbf{X}\boldsymbol{\beta}_d+\boldsymbol{\varepsilon}$ となる．ただし $\mathbf{Y}_d=\mathbf{Y}-\mathbf{X}\boldsymbol{\beta}_0$．この場合には，推定量は $\mathbf{Y}_d'\mathbf{A}\mathbf{Y}_d$ であり，$\mathbf{A}\mathbf{X}=0$ のときこれはすべての $\boldsymbol{\beta}_0$ にたいして $\mathbf{Y}'\mathbf{A}\mathbf{Y}$ と同じである．この必要十分条件は母数 $\boldsymbol{\beta}$ の転移にたいする不変性の条件と呼ばれる．さてこのような推定量にたいして

$$E(\mathbf{Y}'\mathbf{A}\mathbf{Y})=\mathrm{Tr}\,\mathbf{A}\,D(\boldsymbol{\varepsilon})=\sum\gamma_i^2\,\mathrm{Tr}\,\mathbf{A}\mathbf{T}_i=\sum\gamma_i^2q_i$$
$$\Rightarrow \mathrm{Tr}\,\mathbf{A}\mathbf{T}_i=q_i,\qquad i=1,\cdots,k. \tag{4j.2.2}$$

これは不変性のもとでの不偏性の条件である．

（b）最小ノルム　もし仮説的変量 $\boldsymbol{\eta}_i$ がわかっていたとするならば，γ_i^2 の自然な推定量は $\boldsymbol{\eta}_i'\boldsymbol{\eta}_i/n_i$ であり，したがって $\sum q_i\gamma_i^2$ の推定量は

$$(q_1/n_1)\boldsymbol{\eta}_1'\boldsymbol{\eta}_1+\cdots+(q_k/n_k)\boldsymbol{\eta}_k'\boldsymbol{\eta}_k=\boldsymbol{\eta}'\boldsymbol{\Delta}\boldsymbol{\eta}\quad（とおく） \tag{4j.2.3}$$

である．ここで，$\boldsymbol{\Delta}$ は適当に定義された対角行列であり，$\boldsymbol{\eta}'=(\boldsymbol{\eta}_1'\vdots\cdots\vdots\boldsymbol{\eta}_k')$．しかし提起される推定量は，$\mathbf{A}\mathbf{X}=0$ であるから

$$\mathbf{Y}'\mathbf{A}\mathbf{Y}=\boldsymbol{\varepsilon}'\mathbf{A}\boldsymbol{\varepsilon}=\boldsymbol{\eta}'\mathbf{W}'\mathbf{A}\mathbf{W}\boldsymbol{\eta} \tag{4j.2.4}$$

である．ただし $\mathbf{W}=(\mathbf{W}_1\vdots\cdots\vdots\mathbf{W}_k)$．(4j.2.3) と (4j.2.4) との間の差は $\boldsymbol{\eta}'(\mathbf{W}'\mathbf{A}\mathbf{W}-\boldsymbol{\Delta})\boldsymbol{\eta}$ であり，これは，

$$\|\mathbf{W}'\mathbf{A}\mathbf{W}-\boldsymbol{\Delta}\| \tag{4j.2.5}$$

を最小にすることにより，ある意味において小さくすることができる．ここでノルムは適当に選ばれる．

$\mathbf{A}\mathbf{X}=0$ と $\mathrm{Tr}\,\mathbf{A}\mathbf{T}_i=q_i$，$i=1,\cdots,k$ という条件のもとで $\|\mathbf{W}'\mathbf{A}\mathbf{W}-\boldsymbol{\Delta}\|$ を最小にするような $\mathbf{A}$ をもつ推定量 $\mathbf{Y}'\mathbf{A}\mathbf{Y}$ は，母数の関数 $\sum q_i\gamma_i^2$ の MINQUE（最小ノルム2次不偏推定量 Minimum Norm Quadratic Unbiased Estimator）と呼ばれる．

文献（Rao (1970a, 1971c, 1971d, 1972b)）においては，MINQUE はまた，不変性を除き不偏性だけを保証するために，$\mathbf{A}\mathbf{X}=0$ の代りに条件 $\mathbf{X}'\mathbf{A}\mathbf{X}=0$ で求められている．これの区別は不変性をもつまたもたない MINQUE としてなされる．しかしここの議論では不変性の条件を課そう．条件 $\mathbf{X}'\mathbf{A}\mathbf{X}=0$ は，分散成分の推定にたいして，Focke and Dewess (1972) により最初に考えられた．

4j.3　ユークリッド・ノルムのもとでの計算

もしユークリッド・ノルム（すなわち，$\|\mathbf{B}\|^2=\mathbf{B}$ のすべての要素の平方和）を選ぶならば，

$$\|\mathbf{W}'\mathbf{A}\mathbf{W}-\boldsymbol{\Delta}\|^2=\mathrm{Tr}\,\mathbf{W}'\mathbf{A}\mathbf{W}\mathbf{W}'\mathbf{A}\mathbf{W}-2\mathrm{Tr}\,\mathbf{W}'\mathbf{A}\mathbf{W}\boldsymbol{\Delta}+\mathrm{Tr}\,\boldsymbol{\Delta}\boldsymbol{\Delta}$$
$$=\mathrm{Tr}\,\mathbf{A}\mathbf{W}\mathbf{W}'\mathbf{A}\mathbf{W}\mathbf{W}'-\mathrm{Tr}\,\boldsymbol{\Delta}\boldsymbol{\Delta}=\mathrm{Tr}\,\mathbf{A}\mathbf{T}\mathbf{A}\mathbf{T}-\mathrm{Tr}\,\boldsymbol{\Delta}\boldsymbol{\Delta}.$$

ここで $\mathbf{T}=\sum\mathbf{T}_i=\sum\alpha_i^2\mathbf{V}_i$ であり

$$\mathrm{Tr}\,\mathbf{W}'\mathbf{A}\mathbf{W}\boldsymbol{\Delta}=\mathrm{Tr}\,\mathbf{A}\mathbf{W}\boldsymbol{\Delta}\mathbf{W}'=\mathrm{Tr}\sum\frac{q_i}{n_i}\mathbf{A}\mathbf{T}_i=\mathrm{Tr}\,\boldsymbol{\Delta}\boldsymbol{\Delta}$$

という性質を使った．かくして，ユークリッド・ノルムの場合には，問題は，条件 $\mathbf{A}\mathbf{X}=0$，$\mathrm{Tr}\,\mathbf{A}\mathbf{T}_i=q_i$，$i=1,\cdots,k$ のもとで $\mathrm{Tr}\,\mathbf{A}\mathbf{T}\mathbf{A}\mathbf{T}$ を最小にするような $\mathbf{A}$ を決定する問題に帰着される．

この代数的問題はすでに［(iii)，1f.3］に解かれている．そこでは，最小値は

$$\mathbf{A}=\sum\lambda_i\mathbf{R}\mathbf{T}_i\mathbf{R},\qquad \mathbf{R}=\mathbf{T}^{-1}-\mathbf{T}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{T}^{-1}\mathbf{X})^{-}\mathbf{X}'\mathbf{T}^{-1} \tag{4j.3.1}$$

で到達されることが示されている．ただし λ_i は方程式

$$\sum_{i=1}^{k}\lambda_i\mathrm{Tr}\,\mathbf{R}\mathbf{T}_i\mathbf{R}\mathbf{T}_j=q_j,\qquad j=1,\cdots,k \tag{4j.3.2}$$

を満足する．したがって $\sum q_i\gamma_i^2$ の MINQUE は $\boldsymbol{\lambda}'\mathbf{Q}$ である．ただし $\boldsymbol{\lambda}'=(\lambda_1,\cdots,\lambda_k)$ は (4j.3.2) の解であり $\mathbf{Q}'=(Q_1,\cdots,Q_k)$，$Q_i=\mathbf{Y}'\mathbf{R}\mathbf{T}_i\mathbf{R}\mathbf{Y}$．方程式 (4j.3.2) を $\mathbf{S}\boldsymbol{\lambda}=\mathbf{q}$ の形に書いてもよく，その場合には $\boldsymbol{\lambda}=\mathbf{S}^{-}\mathbf{q}$ が1つの解であり，そして

$$\lambda'\mathbf{Q} = \mathbf{q}'(\mathbf{S}^-)'\mathbf{Q} = \mathbf{q}'\mathbf{S}^-\mathbf{Q} = \mathbf{q}'\hat{\gamma}$$

である．ここで $\hat{\gamma}$ は $\mathbf{S}\gamma = \mathbf{Q}$ の1つの解である．

かくして計算手順は，まず方程式 $\mathbf{S}\gamma = \mathbf{Q}$ をつくり，そして1つの解 $\hat{\gamma} = \mathbf{S}^-\mathbf{Q}$ を求めることとなる．ただし $\mathbf{S}$ の (i, j) 要素は $\mathrm{Tr}\,\mathbf{R}\,\mathbf{T}_i\mathbf{R}\,\mathbf{T}_j$ であり，$\mathbf{Q}$ の第 i 要素は $\mathbf{Y}'\mathbf{R}\mathbf{T}_i\mathbf{R}\mathbf{Y}$ である．もとの分散成分にたいする解は，任意の推定可能な関数 $\mathbf{p}'\sigma$ の推定値は $\mathbf{p}'\hat{\sigma}$ であるという意味において，

$$\hat{\sigma} = (\hat{\sigma}_1^2, \cdots, \hat{\sigma}_k^2) = (\alpha_1^2\hat{\gamma}_1^2, \cdots, \alpha_k^2\hat{\gamma}_k^2)$$

である．［$\mathbf{p}'\sigma$ または同値な表現として $\mathbf{q}'\gamma$ は，$\mathbf{AX} = \mathbf{0}$ かつ $\mathrm{Tr}\,\mathbf{AT}_i = q_i$，$i = 1, \cdots, k$ であるような行列 $\mathbf{A}$ が存在するとき推定可能であると呼ばれる．］

$\mathbf{Y}$ が n 変量正規分布をし（これは，(4j. 1. 1) におけるすべての仮説的な変量が1変量正規分布をするということと同値である），真の分散成分が $\alpha_1^2, \cdots, \alpha_k^2$ であるときには，$\mathrm{Tr}\,\mathbf{ATAT}$ は $V(\mathbf{Y}'\mathbf{AY})$ に比例することが確かめられる．したがって仮説的な変量の正規分布のもとでは，MINQUE は局所的最小分散不偏2次推定量である．

4k　線形模型における偏りのある推定

4k. 1　最良線形推定 BLE (best linear estimation)

GGM 模型 $(\mathbf{Y}, \mathbf{X}\beta, \sigma^2\mathbf{G})$ を考え，$\mathbf{L}'\mathbf{Y}$ を，考えられた線形な母数の関数 $\mathbf{p}'\beta$ の1つの推定量とする．$\mathbf{L}'\mathbf{Y}$ の平均2乗誤差 MSE (mean square error) は

$$E(\mathbf{L}'\mathbf{Y} - \mathbf{p}'\beta)^2 = \sigma^2\mathbf{L}'\mathbf{GL} + (\mathbf{X}'\mathbf{L} - \mathbf{p})'\beta\beta'(\mathbf{X}'\mathbf{L} - \mathbf{p}) \qquad (4k. 1. 1)$$

であり，これは未知母数 σ^2 と β の両方を含み，最小にするのに適当な判定基準ではない．したがって3つの可能性がある．

（a）以前の知識をもとにして $\sigma^{-1}\beta$ の事前の値，たとえば $\mathbf{b}$，を選び，(4k. 1. 1) の $\beta\beta'$ に $\sigma^2\mathbf{W} = \sigma^2\mathbf{bb}'$ を代入する．そうすると (4k. 1. 1) は $\sigma^2\mathbf{S}$ の形となる．ただし

$$S = \mathbf{L}'\mathbf{GL} + (\mathbf{X}'\mathbf{L} - \mathbf{p})'\mathbf{W}(\mathbf{X}'\mathbf{L} - \mathbf{p}). \qquad (4k. 1. 2)$$

（b）β を事前の分散共分散 $E(\beta\beta') = \sigma^2\mathbf{W}$ をもつ確率変数と考えるならば，(4k. 1. 1) に β に関しての期待値をもう一度とると，(4k. 1. 1) は $\sigma^2 S$ に帰着される．

（c）(4k. 1. 1) は，1つは分散をあらわし，いま1つは偏りをあらわす2つの部分から成り立っていることがわかる．この場合には，S における $\mathbf{W}$ の選定は分散と比較して偏りに与える相対的な重みをわあらす．したがって推定量の偏りと分散の相対的重要性に関連して $\mathbf{W}$ を適当に選んでよい．

(b) と (c) の場合には，行列 $\mathbf{W}$ は n. n. d. であり　その階数は多分1より大であるう．(i) の結果 (4k. 1. 3) は，$\mathbf{Y} \in \mathscr{M}(\mathbf{G} + \mathbf{XWX}')$ ならば，その階数は何であっても $\mathbf{W}$ の任意の選定にたいして正しい．

かくして判定基準として S を選ぶことにし，命題 (j) は BLE にたいする答を与える．

（j）$\mathbf{p}'\beta$ の BLE は $\mathbf{p}'\tilde{\beta}$ である．ただし

$$\tilde{\beta} = \mathbf{WX}'(\mathbf{G} + \mathbf{XWX}')^-\mathbf{Y}. \qquad (4k. 1. 3)$$

S の最小値は

$$(\mathbf{G} + \mathbf{XWX}')\mathbf{L} = \mathbf{XWp} \quad \text{または} \quad \mathbf{ML} = \mathbf{XWp},\ \mathbf{M} = \mathbf{G} + \mathbf{XWX}'$$

のとき到達されるから，$\mathbf{L} = \mathbf{M}^-\mathbf{XWp}$ は1つの解であり，$\mathbf{p}'\beta$ の BLE は

$$\mathbf{L}'\mathbf{Y} = \mathbf{p}'\mathbf{WX}'\mathbf{M}^-\mathbf{Y} = \mathbf{p}'\tilde{\beta} \qquad (4k. 1. 4)$$

である．$\mathbf{p}'\beta$ が線形不偏推定量をもとうともつまいと BLE は常に存在する．

BLE と BLUE とを比べるために $\mathbf{X}$ は最大階数をもつと仮定しよう．そうすると β は不偏推定量をもち，$R(\mathbf{M}) = R(\mathbf{G} + \mathbf{XWX}') = R(\mathbf{G} \vdots \mathbf{X})$ である．このとき β の BLUE $\hat{\beta}$ は

$$\hat{\beta} = (\mathbf{X}'\mathbf{M}^-\mathbf{X})^{-1}\mathbf{X}'\mathbf{M}^-\mathbf{Y}, \qquad \text{(4i. 4. 3) を用いた} \qquad (4k. 1. 5)$$

であるが，β の BLE は

$$\tilde{\beta} = \mathbf{WX}'\mathbf{M}^-\mathbf{Y} = \mathbf{WX}'\mathbf{M}^-\mathbf{X}\hat{\beta} = \mathbf{T}\hat{\beta} \qquad \text{（とおく）} \qquad (4k. 1. 6)$$

であり，これは $\hat{\beta}$ と $\tilde{\beta}$ との間の関係を与える．$\tilde{\beta}$ の平均2乗誤差（行列）は

$$E(\tilde{\beta} - \beta)(\tilde{\beta} - \beta)' = \mathbf{T}\,E(\hat{\beta} - \beta)(\hat{\beta} - \beta)'\mathbf{T}' + (\mathbf{T} - \mathbf{I})\beta\beta'(\mathbf{T} - \mathbf{I})'$$

$$\mathbf{F} = \mathbf{TDT}' + (\mathbf{T} - \mathbf{I})\beta\beta'(\mathbf{T} - \mathbf{I})' \qquad (4k. 1. 7)$$

である．ここで $\mathbf{D}$ は $\hat{\beta}$ の分散共分散行列であり，β は真の値である．

行列 $\mathbf{D} - \mathbf{TDT}'$ は n. n. d. であるから，$\mathbf{D} - \mathbf{F}$ は β のある範囲にたいして n. n. d. であろう．かくして，もし β が存在すると思われる領域についてある知識があるならば，BLE が BLUE よりも一様に小さい平均2乗誤差をもつことを保証するように $\mathbf{W}$ を適当に選んでもよい．

読者は Hoerl and Kennard (1970a, 1970b) による関連した研究を参照するとよい．(4k. 1. 6) における $\mathbf{T}$ は，$\hat{\beta}$ を零の方向にもっていく一種の収縮因子であり，これは，少なくとも母数空間のある範囲において不偏推定量を改良するのに用いられる普通の技法である，ということに注意せよ．

4k. 2　最良線形偏り最小推定 BLIMBE

母数 β に制約がある場合を含む GGM 模型 $(\mathbf{Y}, \mathbf{X}\beta, \sigma^2\mathbf{G})$ を考えよう．$R(\mathbf{X})$ が最大でないときには，線形な母数の関数のすべてが線形不偏推定量 (LUE) をもつということはない．事実，$\mathbf{p}'\beta$ が LUE をもつための1つの必要十分条件は $\mathbf{p} \in \mathscr{M}(\mathbf{X}')$ である．2つ

の問題をあげる．

（a）　すべての母数の関数が LUE，したがって BLUE をもつように β につけるべき最小の制約は何であるか？

（b）　$\mathbf{p}'\beta$ が LUE をもたないならば，どのような意味においてそれの BLIMBE を見いだすことができるか？

問題 (a) にたいする答は命題 (i) に，(b) にたいする答は命題 (iii) に含まれている．

（i）　$R(\mathbf{X}) = r < m$ とする．ただし m は母数の数である．β に関する最小の制約は2つのちがった形であらわすことができる．

（a）　$\mathbf{R}\beta = \mathbf{c}$，$R(\mathbf{R}) = m - r$ かつ $\mathscr{M}(\mathbf{R}') \cap \mathscr{M}(\mathbf{X}') = \{0\}$．

（b）　$\beta = \mathbf{d} + \mathbf{T}\mathbf{X}'\gamma$，ただし $\mathbf{T}$ は $R(\mathbf{X}\mathbf{T}\mathbf{X}') = R(\mathbf{X})$ であるような任意の行列であり γ は任意の母数である．

最初の制約は明らかであり，2番目のものは最初のものから導くことができる．両方の制約とも β が次元 r の超平面に制限されることを導く．

第2の問題に答えるために，Chipman (1964) による方法に従う．$\mathbf{p}'\beta$ の推定量としての $\mathbf{L}'\mathbf{Y}$ の偏りは $(\mathbf{X}'\mathbf{L} - \mathbf{p})'\beta$ であり，これは，適当なノルムを選んで $\|\mathbf{X}'\mathbf{L} - \mathbf{p}\|$ が最小であるような $\mathbf{L}$ を見いだすことにより小さくされる．

（ii）　$\|\mathbf{X}'\mathbf{L} - \mathbf{p}\|^2 = (\mathbf{X}'\mathbf{L} - \mathbf{p})'\mathbf{B}(\mathbf{X}'\mathbf{L} - \mathbf{p})$ とする．ただし $\mathbf{B}$ は p. d. である．そうすると $\mathbf{p}'\beta$ の1つの線形偏り最小推定 LIMBE (linear minimum bias estimator) は

$$\mathbf{p}'[(\mathbf{X}')^{-}_{l(\mathbf{B})}]'\mathbf{Y} = \mathbf{p}'\mathbf{X}^{-}_{m(\mathbf{B}^{-1})}\mathbf{Y} \tag{4k. 2. 2}$$

である〔式番 (4k. 2. 1) はない〕．

結果 (4k. 2. 2) は，最小2乗 g 逆行列の定義と，それと最小ノルム g 逆行列との同値性から得られる ((1c. 5. 7) 参照)．

LIMBE は一意ではないかも知れない．この場合には，最小分散をもつものを選び，それは最良線形偏り最小推定量 BLIMBE (best linear minimum bias estimator) と呼ばれる．

（iii）　$\mathbf{p}'\beta$ の BLIMBE，すなわち $(\mathbf{X}'\mathbf{L} - \mathbf{p})'\mathbf{B}(\mathbf{X}'\mathbf{L} - \mathbf{p})$ を最小にするような $\mathbf{L}$ のクラスのなかで $\mathbf{L}'\mathbf{G}\mathbf{L}$ が最小であるような線形関数 $\mathbf{L}'\mathbf{Y}$ は $\mathbf{p}'\hat{\beta}$ である．ただし

$$\hat{\beta} = [(\mathbf{X}')^{+}_{\mathbf{BG}}]'\mathbf{Y} \tag{4k. 2. 3}$$

であって，$(\mathbf{X}')^{+}_{\mathbf{BG}}$ は $\mathbf{X}'$ の最小 $\mathbf{G}$ ノルム・$\mathbf{B}$ 最小2乗・g 逆行列である．

結果 (4k. 2. 3) は最小ノルム・最小2乗・g 逆行列の定義 [(iii), **1c. 5**] から得られる．もし $\mathbf{G}$ が最大階数をもつならば，(1c. 5. 9) の結果を用いると

$$\hat{\beta} = [(\mathbf{X}')^{+}_{\mathbf{BG}}]'\mathbf{Y} = \mathbf{X}^{+}_{\mathbf{G}^{-1}\mathbf{B}^{-1}}\mathbf{Y}. \tag{4k. 2. 4}$$

補足と問題

1　線形推定　4a の模型 $(\mathbf{Y}, \mathbf{X}\beta, \sigma^2\mathbf{I})$ と母数の関数 $\mathbf{P}'\beta$ を考える．$\mathbf{P}, \beta, \mathbf{L}, \mathbf{D}, \lambda$ は列ベクトルをあらわすとする．

1.1　$\mathbf{P} = \mathbf{X}'\mathbf{L}$ が解 $\mathbf{L}$ をもつならば，$\mathbf{L}'\mathbf{Y}$ は $\mathbf{P}'\beta$ にたいして不偏であることを示せ．$\mathbf{L}_0$ を $\mathscr{M}(\mathbf{X})$ に属する解とする（それは存在し一意である）．そうすると，任意の他の解にたいして $V(\mathbf{L}_0'\mathbf{Y}) \leqslant V(\mathbf{L}'\mathbf{Y})$．[13ページ，問題 **16**，**17** を参照]．

1.2　$\mathbf{P}'\beta$ が推定可能であって，それの l. s. e. の分散の $\mathbf{P}'\mathbf{P}$ にたいする比が最大（または最小）であるような $\mathbf{P}$ を求めよ．

1.3　一般に，線形関数 $\mathbf{L}'\mathbf{Y}$ が $E(\mathbf{L}'\mathbf{Y})$ の不偏推定量として最小分散をもつ必要十分条件は $\mathrm{cov}(\mathbf{L}'\mathbf{Y}, \mathbf{D}'\mathbf{Y}) = 0$，すなわち，$E(\mathbf{D}'\mathbf{Y}) = 0$ つまり $\mathbf{D}'\mathbf{X} = 0$ であるようなすべての $\mathbf{D}$ にたいして $\mathbf{L}'\mathbf{D} = 0$，であることを示せ．

1.4　期待値0をもつ独立な $\mathbf{Y}$ の線形関数の数は $n - r$ であることを示せ．ただし $r = R(\mathbf{X})$．

1.5　問題 **1.3**，**1.4** の結果から，最小分散推定量は $\lambda'\mathbf{X}'\mathbf{Y}$ の形であることを示せ．またその逆を示せ．

1.6　最小2乗法の理論の模型 $(\mathbf{Y}, \mathbf{X}\beta, \sigma^2\mathbf{I})$ において，1つの推定量の線形性と不偏性の条件は，未知母数は範囲に制限がないと仮定すると，線形性と有界な平均2乗誤差の条件と同値である（最小2乗推定量を導く他の条件にたいする問題4を参照）．

2　推定量の結合

2.1　$t_1, \cdots, t_k$ は（1次元）母数 θ の不偏推定量で $\mathrm{cov}(t_i, t_j) = \sigma_{ij}$ とするとき，θ にたいして不偏でありかつ最小分散をもつ $t_1, \cdots, t_k$ の線形関数を求めよ．

2.2　$t_1, \cdots, t_k$ は分散 $\sigma_1^2, \cdots, \sigma_k^2$ をもつ θ の無相関な推定量とする．$\sum c_i t_i$，$(\sum c_i = 1)$ は，$c_i = \sigma_i^{-2}/\sum \sigma_i^{-2}$ のとき最小分散をもつことを示せ．到達される最小分散を求めよ．

2.3　$t_1, \cdots, t_k$ が母数 $\theta_1, \cdots, \theta_k$ の不偏最小分散推定量であるならば，$c_1 t_1 + \cdots + c_k t_k$ は $c_1\theta_1 + \cdots + c_k\theta_k$ の不偏最小分散推定量であることを示せ．

3　$E(y_i) = x_{i1}\beta_1 + \cdots + x_{im}\beta_m$ で $\mathrm{cov}(y_i, y_j) = \rho\sigma^2$，$V(y_i) = \sigma^2$ であるような観測値 $(y_1, \cdots, y_n) = \mathbf{Y}'$ を考える．$z_1 = (y_1 + \cdots + y_n)/\sqrt{n}$ であるような $\mathbf{Y}$ から $\mathbf{Z}$ への直交変換をつくる．$\mathbf{Z}_2' = (z_2, \cdots, z_n)$ は無相関で同じ分散 $\sigma^2(1 - \rho)$ をもち $E(\mathbf{Z}_2) = \mathbf{U}\beta$ であることを示せ．ただし $\mathbf{U}$ は $\mathbf{X}$ と変換の特定の行列とに関係する．ρ と σ^2 が未知であるときには，母数の関数の不偏最小分散線形推定量は模型 $[\mathbf{Z}_2, \mathbf{U}\beta, \sigma^2(1 - \rho)\mathbf{I}]$ のもとでの l. s. 推定量であることを示せ．

4　模型 $(\mathbf{Y}, \mathbf{X}\beta, \sigma^2\mathbf{I})$ において，直交変換 $\mathbf{Z}_1 = \mathbf{B}\mathbf{Y}$，$\mathbf{Z}_2 = \mathbf{D}\mathbf{Y}$ [(i), **3b. 5** 参照] をしたとする．ただし $E(\mathbf{Z}_2) \equiv 0$，$R(\mathbf{D}) = n - r$，$E(\mathbf{Z}_1) = \mathbf{B}\mathbf{X}\beta \neq 0$ であって $\mathbf{Z}_1, \mathbf{Z}_2$ のすべての成分は無相関である．$E(\mathbf{Z}_2) \equiv 0$ で $\mathrm{cov}(\mathbf{Z}_1, \mathbf{Z}_2) = 0$ であるから，$\mathbf{Z}_2$ は β にたいしては不必要であるということに同意するであろう．$f(\mathbf{Z}_1)$ を，母数の関数 $\mathbf{P}'\beta$ の一致推定量を与える $\mathbf{Z}_1$ の任意の一般の関数であるとする，すなわち $f(\mathbf{B}\mathbf{X}\beta) \equiv \mathbf{P}'\beta$，ただし，$\mathbf{B}\mathbf{X}\beta$ は $\mathbf{Z}_1$ の期待値または観測値に誤差のないときには $\mathbf{Z}_1$ の値である．$f(\mathbf{Z}_1)$ は $\mathbf{Z}_1$ に関して線形であり，また l. s. 推定量であることを示せ．[かくして，不偏性，推定量の線形性，または最小分散性に訴えることなしに，l. s. 推定の論理的な正当性を得た．観測値 $\mathbf{Y}$ における余分なものは，一般に，精度の推定値を得るためにわざと導入される．かくし

て，$\mathbf{Z}_2$ の成分は，それらだけでは β についての情報を与えないけれども，未知母数 β の推定量の精度の推定を可能にする．]

5　秤量計画　個々の重さを調べたい m 個の品物があるとする．1つの方法は各品物を r 回計り，重さの推定値として平均値をとることである．この手順は全体で mr 回の測定を必要とし，各推定された重さの精度は σ^2/r である．ここで σ^2 は個々の測定における誤差の分散である．もう1つの方法は品物を組み合わせて計ることである．各操作は，いくつかの品物を天びんの1つの皿に，残りのものを他の皿にのせ，そして釣合のとれるように分銅をのせることである．これより，型

$$x_1\omega_1 + \cdots + x_m\omega_m = y$$

の観測方程式を得る．ただし，$\omega_1, \cdots, \omega_m$ は品物の仮説の重さであり，i 番目の品物が使われないか，左の皿にのせられたか，右の皿にのせられたかに従って $x_i = 0, 1, -1$ であり，y は釣合のため必要であった重さである．左右の皿にたいする品物の異なった組合せで示される n 回の操作は n 個の観測方程式を導き，これから未知の重さ $\omega_1, \cdots, \omega_m$ は最小2乗法により推定される．観測方程式の $n \times m$ の計画行列 $\mathbf{X}$ は，各要素が $0, 1, -1$ のどれかである特別な型である．問題は，個々の推定値の精度ができるだけよくなるように $\mathbf{X}$ を選ぶことである．

5.1　**4a.12** の結果を用いて，最大の精度は，$\mathbf{X}$ の各要素が1か -1 であって $\mathbf{X}$ の列が直交するとき到達されることを示せ．

5.2　**5.1** の最適な計画にたいして，各推定値の精度は n 回の測定にたいして σ^2/n である．各品物を個別に計ることによって同じ精度を得るためには，要求される測定回数は，mn であることを示せ．

5.3　次の観測方程式（偏りは無視する）から4個の品物の重さを推定せよ．

ω_1	ω_2	ω_3	ω_4	重さ
1	1	1	1	20.2
1	−1	1	−1	8.1
1	1	−1	−1	9.7
1	−1	−1	1	1.9
1	1	1	1	19.9
1	−1	1	−1	8.3
1	1	−1	−1	10.2
1	−1	−1	1	1.8

推定値の分散共分散行列と各測定値の分散 σ^2 の最小2乗推定値を求めよ．

5.4　次の秤量計画にたいして，ω_2, ω_3 の個々の推定値は求められないが，それらの和は推定可能であることを示せ．正規方程式の行列の g 逆行列を求め，線形関数 $\omega_2 + \omega_3$ とこの推定値の標準誤差とを推定せよ．

ω_1	ω_2	ω_3	重さ
1	·	·	5.1
·	1	1	8.2
1	−1	−1	4.9
−1	1	1	3.1

6　$a_1 + b_1(x - \bar{x}_1)$，$a_2 + b_2(x - \bar{x}_2)$ を，大きさ n_1, n_2 の独立な標本からの2つの推定回帰関数であるとする．さらに $S_{yy}^{(1)}, S_{xy}^{(1)}, S_{xx}^{(1)}$ を第1の標本の修正された平方和，修正された積和であるとし，指数 (2) で第2の標本のそれらをあらわすとする．真の回帰関数が交わるところの x の値 ξ の信頼区間を求めよ．

ヒント：$z = a_1 + b_1(\xi - \bar{x}_1) - a_2 - b_2(\xi - \bar{x}_2)$ とすると，$E(z) = 0$，$V(z) = c\sigma^2$，ただし

$$c = \frac{1}{n_1} + \frac{(\xi - \bar{x}_1)^2}{S_{xx}^{(1)}} + \frac{1}{n_2} + \frac{(\xi - \bar{x}_2)^2}{S_{xx}^{(2)}}.$$

$1-\alpha$ 信頼域は ξ に関する2次式

$$\frac{z^2}{c\hat{\sigma}^2} = F_\alpha(1, n_1 + n_2 - 4)$$

を解くことにより得られる．ただし

$$\hat{\sigma}^2(n_1 + n_2 - 4) = S_{yy}^{(1)} - \frac{[S_{xy}^{(1)}]^2}{S_{xx}^{(1)}} + S_{yy}^{(2)} - \frac{[S_{xy}^{(2)}]^2}{S_{xx}^{(2)}}.$$

7　1つの確率変数は2つの変数，その1つは与えられた変数の集合と重相関1をもち（完全に決定される），いま1つは与えられた変数のおのおのと無相関である，の和として書くことができることを示せ．

8　(ρ_{ij}) を p 変量 $x_1, \cdots, x_p$ の相関行列とする．**4g.1** で与えた偏・重相関係数の定義から，次のことを示せ．

(a)　$1 - \rho_{1(2\cdots p)}^2 = (1 - \rho_{12}^2)(1 - \rho_{13\cdot 2}^2)\cdots(1 - \rho_{1p\cdot(2\cdots p-1)}^2)$

(b)　$\rho_{12\cdot 3} = (\rho_{12} - \rho_{13}\rho_{23})/(1 - \rho_{13}^2)^{\frac{1}{2}}(1 - \rho_{23}^2)^{\frac{1}{2}}$

(c)　$\rho_{12} = (\rho_{12\cdot 3} + \rho_{13\cdot 2}\rho_{23\cdot 1})/(1 - \rho_{13\cdot 2}^2)^{\frac{1}{2}}(1 - \rho_{23\cdot 1}^2)^{\frac{1}{2}}$

(d)　$\rho_{12\cdot(34\cdots p-1)} = \dfrac{\rho_{12\cdot(3\cdots p)} + \rho_{1p\cdot(2\cdots p-1)}\rho_{2p\cdot(13\cdots p-1)}}{(1 - \rho_{1p\cdot(2\cdots p-1)}^2)^{\frac{1}{2}}(1 - \rho_{2p\cdot(13\cdots p-1)}^2)^{\frac{1}{2}}}$

(e)　$\rho_{12\cdot(34\cdots p)} = \dfrac{\rho_{12\cdot(3\cdots p-1)} - \rho_{1p\cdot(3\cdots p-1)}\rho_{2p\cdot(3\cdots p-1)}}{(1 - \rho_{1p\cdot(3\cdots p-1)}^2)^{\frac{1}{2}}(1 - \rho_{2p\cdot(3\cdots p-1)}^2)^{\frac{1}{2}}}$

9　$\rho_{1(2\cdot 34\cdots p)}$ を，x_1 と，$x_3, \cdots, x_p$ にたいする回帰を除去した x_2 の残差との間の相関と定義する．この係数は部分相関と呼ばれる．偏相関とちがって記号1,2に関して対称ではない．次のことを示せ．

(a)　$\rho_{1(2\cdot 34\cdots p)}^2 \leqq \rho_{12\cdot(3\cdots p)}^2$

(b)　$\rho_{1(234\cdots p)}^2 = \rho_{1p}^2 + \rho_{1(p-1\cdot p)}^2 + \cdots + \rho_{1(2\cdot 34\cdots p)}^2$

10　**2つの平行な回帰直線の間の水平距離の推定（フィーラーの定理）**　変量の組 (y, x) についての大きさ n_1 の標本にたいし，$\bar{y}^{(1)}, \bar{x}^{(1)}$ を平均値，$S_{yy}^{(1)}, S_{xy}^{(1)}, S_{xx}^{(1)}$ を平方和，積和とする．また指数 (2) でもって，大きさ n_2 の独立な標本からのそれらの値とする．もし真の回帰は平行であることが与えられているならば，共通な回帰係数の推定値は

$$b = \frac{S_{xy}^{(1)} + S_{xy}^{(2)}}{S_{xx}^{(1)} + S_{xx}^{(2)}} = \frac{S_{xy}}{S_{xx}}$$

であることを示せ．η を2つの回帰直線の間の水平距離とすると，統計量 $d = a_1 - a_2 + b\eta$ は期待値0，分散

$$\sigma^2\left\{\frac{1}{n_1}+\frac{1}{n_2}+\frac{[\bar{x}^{(1)}-\bar{x}^{(2)}-\eta]^2}{S_{xx}}\right\}=\sigma^2 C$$

をもつ，ただし $a_1=\bar{y}^{(1)}-b\bar{x}^{(1)}$，$a_2=\bar{y}^{(2)}-b\bar{x}^{(2)}$．$\sigma^2$ の1つの推定値は

$$\hat{\sigma}^2=S_{yy}^{(1)}+S_{yy}^{(2)}-\frac{S_{xy}^2}{S_{xx}}$$

で，自由度は n_1+n_2-3 である．したがって η の信頼区間は，η に関する2次式

$$\frac{d^2}{\hat{\sigma}^2 C}=F_\alpha(1, n_1+n_2-3)$$

を解くことにより与えられることを示せ．

11　線形模型における推定量の頑健性　$(\mathbf{Y}, \mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{G})$ を考え，$\mathbf{G}=\mathbf{I}$ を仮定したとき $\mathbf{P}'\boldsymbol{\beta}$ の BLUE を $\mathbf{P}'\hat{\boldsymbol{\beta}}$ とする．すべての推定可能な $\mathbf{P}'\boldsymbol{\beta}$ にたいして $\mathbf{P}'\hat{\boldsymbol{\beta}}$ が $(\mathbf{Y}, \mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{G})$ のもとでも BLUE であるならば，次のことが必要かつ十分である．

$$\mathbf{G}=\mathbf{XAX}'+\mathbf{ZBZ}',$$

ただし $\mathbf{A}, \mathbf{B}$ は任意の対称行列であり，$\mathbf{Z}$ は $\mathbf{X}'\mathbf{Z}=\mathbf{0}$ であるような最大階数の行列である［この問題の証明やさらに進んだ議論については Rao (1967c, 1968a, 1971e) を参照］．

参考文献

Aitken, A. C. (1935), On least squares and linear combination of observations, *Proc. Roy. Soc. Edin.* **A**, 55, 42-48.

Bose, R. C. (1950-51), *Least Square Aspects of the Analysis of Variance*, Mimeographed series, no. 9. North Carolina University.

Box, G. E. P. (1954), The exploration and exploitation of response surfaces. Some general considerations and examples, *Biometrics* **10**, 16-60.

Chipman, J. S. (1964), On least squares with insufficient observations, *J. Amer. Statist. Assoc.* **59**, 1078-1111.

Cochran, W. G. (1951), Improvement by means of selection, *Proc.* (second) *Berkeley Symp. on Math. Statist. Prob.*, 449-470.

Cochran, W. G. and G. M. Cox, (1957), *Experimental Designs* (second edition), Wiley, New York.

Fisher, R. A. (1925), *Statistical Methods for Research Workers* (first edition 1925, eleventh edition 1950), Oliver and Boyd, Edinburgh.

Focke, J. and Dewess, G. (1972), Über die Shätzmetnode MINQUE von C. R. Rao und ihre verallgemeinerung, *Math. Operationsforsch. u. Statist.* **3**, 129-144.

Gauss, K. F. (1809), *Werke* **4**, 1-93, Gottingen.

Goldman, A. J. and Zelen, M. (1964), Weak generalized inverses and minimum variance linear unbiased estimation, *J. of Research Nat. Bureau of Standards* **68B**, 151-172.

Hoerl, A. E. and Kennard, R. W. (1970a), Ridge regression : Biased estimation for nonorthogonal problems, *Technometrics* **12**, 55-68.

Hoerl, A. E. and Kennard, R. W. (1970b), Ridge regression : Application to nonorthogonal problems, *Technometrics* **12**, 69-82.

Hooke, B. G. E. (1926), A third study of the English skull with special reference to the Farrington Street crania, *Biometrika* **18**, 1-55.

Hsu, P. L. (1958), On the best unbiased quadratic estimate of variance, *Statist. Res. Mem.* **2**, 91-104.

Kempthorne, O. (1952), *The Design and Analysis of Experiments*, Wiley, New York.

Kempthorne, O. and Nordskog, A. W. (1959), Restricted selection indices, *Biometrics*, **15**, 10-19.

Linnik, Yu. V. (1961), *Method of Least Squares and Principles of the Theory of Observations* (Translated from Russian), Pergamon Press, London.

Markoff, A. A. (1900), *Wahrscheinlichkeitsrechnung*, Tebner, Leipzig.

Miliiken, George A. and Graybill, Franklin, A. (1970), Extensions of the general linear hypothesis model, *J. Amer. Statist. Assoc.* **65**, 797-807.

Neyman, J. and F. N. David (1938), Extension of the Markoff theorem on least squares, *Stst. Res. Mem.* **2**, 105-116.

Parzen, E. (1961), An approach to time series analysis, *Ann. Math. Statist.* **32**, 951-989.

Pearson, K. (1899), Mathematical contributions to the theory of evolution V. On the reconstruction of the stature of prehistoric races, *Philos. Trans. Roy. Soc.* **A192**, 169-244.

Plackett, R. L. (1960), *Principles of Regression Analysis*, Clarendon Press, Oxford.

Scheffe, Henry (1959), *The Analysis of Variance*. Wiley, New York.

Tukey, John (1949), One degree of freedom for non-additivity, *Biometrics* **5**, 232-242.

Williams, E. J. (1959), *Regression Analysis*, Wiley, New York.

Wishart, J. (1938), Growth rate determinations in nutrition studies with the bacon pig and their analysis, *Biometrika* **30**, 16-28.

Yates, F. (1949), *Sampling Methods for Censuses and Surveys*, Charles Griffin, London.

Zyskind, G. and Martin, F. B. (1969), On best linear estimation and a general Gauss-Markoff theorem in linear models with arbitrary non-negative covariance structure, *SIAM, J. Appl. Math.* **17**, 1190-1202.

統計数値表

Biometrika tables (1956, 4th edition), *Biometrika Tables for Statisticians* Vol. 1. By E. S. Pearson and H. O. Hartley, Cambridge University Press, Cambridge.

R. M. M. tables (1965), *Tables and Formulae in Statistical Work*. By C. R. Rao, A. Matthai, and S. K. Mitra, Statistical Publishing Society, Calcutta.

F. Y. tables (1963) (6th edition). *Statistical Tables For Biological, Agricultural and Medical Research*. By R. A. Fisher and F. Yates, Oliver and Boyd, London.